



A University of Sussex PhD thesis

Available online via Sussex Research Online:

<http://sro.sussex.ac.uk/>

This thesis is protected by copyright which belongs to the author.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Please visit Sussex Research Online for more information and further details

Empirical Essays on Development Economics

Jorge García Hombrados

Submitted for the degree of Doctor of Philosophy in Economics

University of Sussex

July 2017

Declaration

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree.

Signature:

Jorge García Hombrados

UNIVERSITY OF SUSSEX

JORGE GARCIA HOMBRADOS
DOCTOR OF PHILOSOPHY IN ECONOMICS

EMPIRICAL ESSAYS ON DEVELOPMENT ECONOMICSSUMMARY

This thesis investigates empirically three questions of key relevance for the life of disadvantaged people in developing countries.

Using a sample of Ethiopian women and a regression discontinuity design exploiting age discontinuities in exposure to a law that raised the legal age of marriage for women, the first chapter documents for the first time (a) the effect of increasing the legal age of marriage for women on infant mortality and (b) the causal effect of early cohabitation on infant mortality. The analysis shows that, even though it was not perfectly enforced, the law that raised the legal age of marriage had a large effect on the infant mortality of the first born child. Furthermore, the estimates suggest that the effect of a one-year delay in women's age at cohabitation on the infant mortality of the first born is comparable to the joint effect on child mortality of measles, BCG, DPT, Polio and Maternal Tetanus vaccinations.

Using longitudinal data from northern Ghana, the second chapter shows that parents allocate more schooling to children that are more cognitively able. These results provide evidence for the main prediction of the model of intra-household allocation of resources developed in [Becker \(1981\)](#), which concludes that parents allocate human capital investments reinforcing cognitive differences between siblings.

The third chapter uses the 8.8 Richter magnitude earthquake that struck Chile in February 2010 as a case study and employs a difference in difference strategy to investigate whether natural disasters have lasting effects on property crime. The results show that the earthquake reduced the prevalence of property crime the year of the earthquake and that this effect remained stable over the 4 post-earthquake years studied. The lasting drop in crime rates in affected areas seems to be linked to the earthquake strengthening community life in these municipalities.

Acknowledgements

I owe Edoardo Masset and Andy McKay my deepest gratitude for supervising me through the research conducted for this thesis. Their guidance, support and advice have been crucial for the elaboration of this document. The data used in chapter 2 was facilitated by Edoardo Masset, principal investigator of the Impact Evaluation of the Millennium Villages Project in Northern Ghana. Tsegay Tekleselassie, Nemera Gebeyehu Mamo and Sadri Saieb helped me to gain access and translate regional laws from Ethiopia.

I also want to expand my gratitude to other members of the Sussex faculty. Sam Marden, Richard Dickens and Rocco D’Este provided priceless comments to at least one of the thesis chapters. James Fenske also made valuable comments to an early version of the first chapter. My fellows PhD candidates at the Department of Economics provided a stimulating environment for the discussion and development of research.

This thesis benefited from the financial support of the Economic and Social Research Council (ESRC).

Coni, Inés, Laura and Miguel are my main sources of motivation and support. I would have not been able to overcome the difficulties faced in the last years without their wise advice, unlimited patience and unconditional love. From my father Peregrino, I learned the values of constancy and hard work that enable the completion of a Ph.D. thesis. The work that leads to this thesis pretends to be a tribute to his memory and values. *Vista, suerte y al toro.*

Contents

List of Tables	ix
List of Figures	xi
Introduction	1
1 Child Marriage and Infant Mortality: Evidence from Ethiopia	9
1.1 Motivation	9
1.2 Related Literature	12
1.3 Child Marriage in Ethiopia and the Revised Family Code	14
1.4 Identification Strategy	18
1.5 Data and Descriptive Statistics	22
1.6 Results	26
1.6.1 The Effect of the RFC on the Age at Cohabitation	26
1.6.2 The Effect of the RFC on Infant Mortality	30
1.6.3 The Effect of Women's Age at Cohabitation on Infant Mortality	32
1.6.4 Robustness Checks	32
1.7 Mechanisms	42
1.8 Conclusions	45
Appendix 1.A McCrary Test: Density of the Forcing Variable at the Cut-Off	47
Appendix 1.B Additional Graphs	48
Appendix 1.C Sample of Women Still Cohabiting with the First Partner	53
2 Cognitive Skills and Intra-Household Allocation of Schooling	54
2.1 Introduction	54
2.2 Conceptual framework: Becker (1981)	56
2.3 Related Literature	58

2.4	Education in Ghana	60
2.5	Data	61
2.6	Empirical Strategy	68
2.7	Results	70
2.8	Heterogeneous Effects	74
2.8.1	Gender	74
2.8.2	Household Consumption	77
2.8.3	Polygyny	79
2.8.4	Household Size	80
2.9	Conclusions	81
	Appendix 2.A Appendix	83
3	The Lasting Effects of Natural Disasters on Property Crime: Evidence from the 2010 Chilean Earthquake	85
3.1	Motivation	85
3.2	Conceptual Framework	87
3.3	Related Literature	89
3.4	The Context	90
3.5	Data	92
3.6	Empirical Strategy	98
3.7	Results	102
3.8	Additional Analysis: Known to the Police Crime Data	106
3.8.1	The Effect of the Earthquake on Known to the Police Property Crime	106
3.8.2	Crime and Punishment in the Aftermath of the Earthquake	112
3.9	Analysis of Mechanisms	114
3.10	Conclusions	121
	Appendix 3.A Robustness Checks: Using All Observations and Heterogeneity of Effects	123
	Appendix 3.B Earthquake Intensity Scales	128
	Appendix 3.C Maps: Treatment, Control and Excluded Municipalities under the Use of Different Distance Thresholds	130

Appendix 3.D Reporting Rate for Different Types of Crime (ENUSC Data):	
Analysis at the Regional Level	132
Appendix 3.E Additional Graphs	134
Conclusions	135
Bibliography	138

List of Tables

1.1	Summary statistics: Women that ever cohabited and ever bore a child in the regions included in the study.	24
1.2	Non-parametric methods: RFC, age at first cohabitation and infant mortality.	28
1.3	Parametric methods using different time windows: RFC, age at first cohabitation and infant mortality.	29
1.4	RFC and different outcomes.	33
1.5	Robustness checks Infant Mortality: Placebo analyses.	35
1.6	Analysis of mechanisms	44
1.7	Only women still cohabiting with first cohabit.	53
2.1	Summary statistics: Individual characteristics of children aged 5-18 in 2012.	64
2.2	Summary statistics: Household characteristics of children aged 5-18 years in 2012.	67
2.3	First stage (Probit coefficients): Participation of children in cognitive skills tests (children 5-18 in 2012).	71
2.4	Child cognitive skills and years of schooling 2012-2015 (children 5-18 in 2012).	72
2.5	Child cognitive skills and years of schooling 2012-2015 (Children 5-12 in 2012).	73
2.6	Effect by gender: child cognitive skills and years of schooling 2012-2015 (children 5-18 in 2012).	76
2.7	Effect by household characteristics: child cognitive skills and years of schooling 2012-2015 (children 5-18 in 2012).	78
2.8	First stage (Probit coefficients): Participation of children 5-18 in 2012 in cognitive skills tests (school attendance).	83
2.9	Child cognitive skills and school attendance 2013-15 (children 5-18).	84

3.1	Fatalities and economic damage of the earthquake/tsunami by region	92
3.2	Descriptive Statistics for variables used in the analysis	96
3.3	Effects of the earthquake on home burglary (ENUSC data): Leads and lags analysis and pooled effects for the period 2007-2013	105
3.4	Effects of earthquake exposure on property crime (2007-2013): SPD data .	109
3.5	Impact estimates (OLS): Short-term effects of the earthquake on different types of property crimes and on individuals apprehended (SPD data)	113
3.6	Effects of the earthquake on social capital and the adoption of individual and community-based measures to prevent crime	117
3.7	The effects of the earthquake on other sociodemographic and economic vari- ables	120
3.8	Effects of the earthquake on home burglary: Different samples, municipality time trends and heterogeneity of effects	123
3.9	Impact estimates (OLS): Short-term effects of the earthquake on different types of property crimes and on individuals apprehended (SPD data)	124
3.10	The effects of the earthquake on other sociodemographic and economic vari- ables	125
3.11	Effect of the earthquake on the probability of reporting a crime to the police and mean reporting rates (regional level analysis)	132

List of Figures

1.1	Child marriage over time in Ethiopia (DHS 2011)	15
1.2	Age at 1st cohabitation (cohorts 15-17 at RFC): Discontinuities at 15 and 18	17
1.3	Age at 1st cohabitation (cohorts 12-14 at RFC): Discontinuities at 15 and 18	17
1.4	Age at cohabitation, age at birth and infant mortality of the first born (LOWESS regressions)	25
1.5	Main analysis: Age at first cohabitation at the cut-off	27
1.6	Main analysis: Infant mortality rate at the cut-off	31
1.7	Placebo variables: ethnicity, religion and gender	36
1.8	Placebo test: Cut-off at 19 years	37
1.9	Placebo test: Discontinuity in other Ethiopian regions	38
1.10	RFC and women characteristics	41
1.11	Density of the forcing variable at the cut-off.	47
1.12	RFC and child marriage: Includes all women (not only those that ever cohabit and gave birth)	48
1.13	Age at first cohabitation and infant mortality of first born: Full sample of women (18-49)	49
1.14	Age at 1st cohabitation by age cohort.	49
1.15	Mechanisms: Fertility, marriage market outcomes and other women outcomes	50
1.16	Mechanisms: Maternity and child health for first born	51
1.17	Age at first cohabitation at the cut-off for the Ethiopian regions that ap- proved RFC between 2000 and 2007	52
2.1	School attendance by age (year 2012)	66
3.1	Predicted intensity: treatment and control areas	95
3.2	Incidence of home burglary over time	100

3.3	Effect of the earthquake on home burglary over time	106
3.4	Incidence of crime over time (SPD data 2007-2013)	110
3.5	Incidence of crime over time (SPD data 2003-2013)	111
3.6	Incidence of crime over time (SPD data 2007-2013)	126
3.7	Incidence of crime over time (SPD data 2003-2013)	127
3.8	Treatment and Control areas	130
3.9	Treatment and Control areas	131
3.10	Evolution of reporting rate by type of crime (ENUSC data)	133
3.11	Effects of the earthquake on home burglary over time (ENUSC data): Dif- ferent thresholds used to construct treatment and control municipalities . .	134

Introduction

Does child marriage affect infant mortality? Do parents in low-income countries allocate schooling correcting or reinforcing cognitive differences between siblings? Do natural disasters affect crime rates in a lasting way? The chapters that formed this thesis address these research questions.

Chapter 1 focuses on examining the inter-generational effects of child marriage in Ethiopia. Deeply rooted in tradition, gender inequality and poverty ([Wahhaj, 2015](#)), child marriage affects more than 700 million women worldwide and 41% of women aged 20-24 in Ethiopia ([UNFPA, 2012](#)). Considered by UNICEF a form of violence against women, several developing countries with high incidence of child marriage aimed to address this problem through ratifying international conventions such as the Maputo Protocol or the Convention on the Elimination of All Forms of Discrimination against Women that promoted the approval and enforcement of minimum-age-of-marriage laws. In parallel with these legal reforms, many of these countries and different international organizations have launched ambitious programmes targeting the socioeconomic and cultural causes of child marriage.

Despite plenty of resources have been spent in fighting child marriage, there is little empirical evidence on the socioeconomic effects of policies aiming to reduce this practice. There is also a lack of rigorous evidence in the literature on the assessment of its causal effects on socioeconomic outcomes. Existing studies rely on the use of age at menarche as an instrumental variable for the age at marriage, an approach that is challenged by some

medical studies.

Using data from Ethiopia, chapter 1 addresses two questions of key importance on the link between child marriage and infant mortality. First, the chapter investigates the effects of a law that raised the minimum age of marriage for women from 15 to 18 years on women's age at cohabitation and on the infant mortality of the first born child. Second, the analysis exploits this policy change to estimate the causal effect of women's age at cohabitation on the probability of infant mortality of the first born child.

The empirical results highlight that exposure to a legal age of marriage at 18 years relative to the possibility of getting legally married at 15, increases by 2 years the mean age at first cohabitation for women and decreases by 7.9 percentage points the probability of infant mortality of the first born child. The reduction in the probability of infant mortality of the first born child caused by a one-year delay in women's age at cohabitation during teenage years is estimated at 3.8 percentage points.

This chapter makes three contributions to the literature. First of all, it provides the first empirical evidence on the socioeconomic effects of minimum-age-of-marriage laws, showing that they can contribute to reducing child marriage and infant mortality even when they are not completely enforced. Second, this is the first study that provides causal estimates on the effect of early cohabitation on infant mortality. Third, this study uses a novel regression discontinuity design that can be used to (a) expand the analysis on the effects of child marriage to other key outcomes and settings where similar laws have been approved and to (b) test the robustness of the results from previous studies that rely on age at menarche as an instrumental variable for age at marriage.

In 1980, the winner of the Nobel Prize in Economics Gary Becker proposes a theoretical model to explain the intra-household allocation of resources based on the different returns to investments on each household member. The model predicts that through allocating more human capital resources to the most endowed children, parents reinforce cognitive

differences between children.

Chapter 2 of the thesis investigates empirically this hypothesis through exploring how between siblings cognitive differences affect the intra-household allocation of schooling. For this, I use child level panel data on education collected for the impact evaluation of the Millennium Villages Project in northern Ghana and the score in three cognitive tests applied with the first round of the survey to every child in the sample aged 5-19 regardless of whether the child has ever attended school.

The empirical analysis reveals that, in line with the prediction of Becker’s model and with previous empirical evidence, parents tend to allocate more schooling to children with better cognitive skills. Interestingly, the evidence provided in this chapter also shows that some characteristics prevalent in sub-Saharan households such as large household size, polygyny or poverty do not seem to affect the role played by cognitive skills in the allocation of schooling across siblings.

The contribution of this chapter to the literature is twofold. On the one hand, the evidence presented adds to the small body of literature that uses cognitive tests to examine empirically the effect of cognitive skills on the intra-household allocation of human capital resources. Second, this is to the best of my knowledge the first study that assesses empirically whether the effect of cognitive skills on the allocation of schooling across siblings depends on household characteristics such as polygyny status, household size or wealth.

Chapter 3 uses data from household victimization surveys and a strong earthquake that affected the centre-south of Chile in 2010 as a case study to explore the lasting effects of natural disasters on property crime. The earthquake was followed by looting and social chaos. However, the evidence also shows that the earthquake strengthened community links in affected communities and the degree in which households committed with the well-being of their neighbours (Grandón et al., 2014). Thus, the earthquake affected ambiguously the benefits and costs of crime and the assessment of its long-term

effects on crime remains an empirical question.

Using a difference in difference strategy, the empirical analysis shows that the incidence of property crime in areas affected by the earthquake fell dramatically the year of the earthquake relative to communities far away from the hypocentre. Interestingly, the difference in the incidence of property crime between earthquake affected and unaffected municipalities remained stable and statistically significant during the 4 post-earthquake years studied. The chapter also presents evidence on the mechanisms concluding that the lasting drop in property crime in municipalities close to the hypocentre was mainly caused by the positive effect of the earthquake on community links and on the adoption of community-based measures to prevent crime by households living in these municipalities.

The contribution of chapter 3 is twofold. First, although previous papers discuss crime in the aftermath of natural disasters, this is the first study testing whether the short-term effects of natural disasters on crime last over longer periods of time. Second, the chapter documents for the first time the role that social capital and informal guardianship played as decisive factors that explain the effect of natural disasters on property crime.

Seemingly unrelated as these three chapters seem to be, there are three common themes among them.

First, the three chapters address questions with important policy implications for the well-being of disadvantaged people in low- and middle-income countries. Child marriage and insufficient schooling investments are widespread problems among the most disadvantaged households in developing countries. For example, [Klugman et al. \(2014\)](#) show that girls from poor households in sub-Saharan Africa are twice more likely to engage in underage marriages than non-poor girls, and the lack of attendance to primary and secondary school in developing countries is strongly associated with household poverty ([UN, 2013](#)). Understanding the causes and consequences of these two phenomena has become a priority for policy makers of international organizations and low- and middle-income

countries that have signed conventions and have dedicated a large amount of resources to promote universal schooling and increase the age at cohabitation for women.

In this context, chapter 1 documents the negative effects of child marriage on infant mortality in Ethiopia, the country with the fifth largest number of women married before the age of 18 years, and shows that raising the legal age of marriage can contribute to decreasing infant mortality. On the other hand, chapter 2 provides evidence on how parents in poor areas in Ghana tend to concentrate schooling investments in the more able siblings rather than spreading the investments among all of them, casting doubts on whether supply-side educational interventions such as reducing class sizes or providing more books could benefit less able children.

Although there has been a rise in the incidence of natural disasters worldwide in the last decade, the most disadvantaged people in poor countries faced most of the negative consequences of them (Yonetani, 2012). Understanding the impacts of natural disasters and minimizing the vulnerability to them have become priorities in the political agenda of countries that are frequently exposed to natural disasters such as Bangladesh, Myanmar or Chile (UNISDR, 2013). Through providing evidence on the lasting effects of natural disasters on crime and on the role of informal guardianship in the aftermath of disasters, chapter 3 contributes to shedding light on the socioeconomic consequences and social dynamics that follow natural disasters in a country that is regularly exposed to geological disasters.

The second common aspect is the central position of social norms in the development problems studied and empirical results obtained in the three chapters.

Mackie et al. (2015) defines social norms as “what people in some group believe to be a typical or appropriate action in the group”. The relevance of social norms as determinants of economic and social outcomes was first highlighted by David Hume in his Treatise of

Human Nature (1739)¹. In the last decades, different studies have examined how social norms can affect different dimensions of individual's well-being. For example, [Duflo and Udry \(2004\)](#) investigate how social norms affect the intra-household allocation of resources and [Mackie et al. \(2015\)](#) argue that social norms contribute to the perpetuation of harmful practices such as female genital mutilation or child marriage. In this sense, an important part of chapter 1 is the assessment of whether a law that changed the legal age of marriage could tackle effectively a social practice deeply enrooted in tradition: child marriage. Equally, chapter 2 investigates whether the social norms that govern the intra-household allocation of resources reinforce cognitive differences between the children of the household.

On the other hand, the evidence presented in chapter 3 suggests that natural disasters are followed by changes in the social norms that function in affected communities, enabling the adoption of community-based crime prevention strategies that could ultimately reduce crime.

The third common theme among the chapters that formed the thesis is methodological. The analyses conducted in them rely on econometric methods that fall within those traditionally used in applied microeconomics for the identification of causal effects and more recently, for the impact evaluation of policies or programmes.

Chapter 1 uses sharp and fuzzy regression discontinuity designs (RDD) to assess the causal link between the legal age of marriage, the prevalence of child marriage and the infant mortality of the first born. Firstly used by [Thistlethwaite and Campbell \(1960\)](#) in the field of educational psychology to investigate the motivational effect of public recognition, RDDs became widespread in economics in the last decade as a credible identification strategy to assess causal effects and to evaluate social programmes.

Fuzzy and sharp RDDs can be used when the probability of treatment assignment for each participant is partially or entirely determined by whether the value of the assignment

¹[Hume \(2000\)](#)

variable for a participant is above or below a certain cut-off point. In chapter 1, I exploit the fact that women younger than 15 years when the law that raised the legal age of marriage was enacted could not get legally married until they turned 18 years old while those girls aged 15 or older when the law was introduced had the opportunity to get legally married before they reached the age of 18. This *age at policy* discontinuity is used to investigate the effect for women of exposure to a legal age of marriage at 18 on the probability of infant mortality of the first born child and then, to calculate the effect of women's age at cohabitation during teenage years on infant mortality of the first born child.

The main advantage of RDDs is the possibility of testing empirically their underlying identification assumptions through examining the density of the forcing variable at the cut-off and investigating discontinuities in placebo variables at the cut-off. On the other hand, the main limitation of this method is that the estimates yielded by RDDs can only be interpreted as local treatment effects at the cut-off. Thus, any generalization of the effects to the whole sample should be conducted with caution.

The second chapter uses a Heckman selection model to account for potential sample selection bias arising from the fact that a substantial amount of eligible children did not take the cognitive tests. Introduced by Heckman (1979) to account for non-random selection of women participating in the labour market, this method was quickly adopted in the economics literature to deal with non-random censoring in the sample. The method is also commonly used in the field of applied microeconomics to investigate causal effects. Some examples are Cueto et al. (2014) and Briggs (2004), where Heckman selection models are used to estimate the causal effects of school quality on educational attainment and of a coaching programme on exam performance; while accounting for non-random selection of participants into school attendance and exam uptake.

The third chapter relies on the use of difference in difference estimations. Firstly

employed in the XIX century by John Snow to assess the effects of contaminated water on cholera incidence in London ([Angrist and Pischke, 2008](#)), the method became very popular in the last two decades. This estimation technique is widely used in the policy evaluation literature when data before and after a policy change or programme are available.

Indeed, difference in difference techniques have been used in seminal papers to address a large variety of crucial research questions such as the causal effect of minimum wage on employment ([Card and Krueger, 1994](#)), the effect of school term length on student performance ([Pischke, 2007](#)), or the effect of employment protection on outsourcing by firms ([Autor, 2003](#)). In chapter 3, I use a difference in difference approach, comparing the evolution of crime rates in areas far away and close to the earthquake hypocentre before and after the strong earthquake that affected Chile in February 2010, to assess the persistent effects of the earthquake on property crime.

Chapter 1

Child Marriage and Infant Mortality: Evidence from Ethiopia

1.1 Motivation

More than 700 million women worldwide first cohabited with a partner before the age of 18, with the vast majority living in developing countries ([UNICEF, 2014](#)). Considered by UNICEF a form of violence against women, decreasing the incidence of early cohabitation, also known as child marriage¹, has become a priority for policy makers of international organizations and developing countries. In the last decades, most of the countries with a high prevalence of this practice have ratified different international agreements such as the CEDAW², CCMMAMRM³ or the Maputo Protocol⁴ that promote the setting and enforcement of minimum-age-of-marriage laws. Furthermore, the fight against child marriage mobilizes a large amount of resources in integrated programmes and national alliances targeting the cultural, social and economic causes of this widespread practice. However, although the overall prevalence of child marriage is declining over time, its eradication is currently far from becoming a reality ([Jensen and Thornton, 2003](#)).

Using women-level data from different Asian and African countries, several studies

¹UNICEF defines child marriage as the formal marriage or unmarried cohabitation before the age of 18 years. Although child marriage affects both girls and boys, the 82% of the children in the world that got married or started cohabiting with a partner before the age of 18 are girls ([UNICEF, 2014](#)).

²Convention on the Elimination of All Forms of Discrimination against Women, 1979.

³Convention on Consent to Marriage, Minimum Age for Marriage, and Registration of Marriage, 1964.

⁴Ratified in 2003.

reveal that child marriage is associated with worse levels of health, education, labour force participation, mortality and participation in household decisions (Jensen and Thornton, 2003; Nguyen and Wodon, 2015; Wodon et al., 2016; Elborgh-Woytek et al., 2013; Wachs, 2008; UNICEF, 2014; Elborgh-Woytek et al., 2007; Solanke, 2015). However, the link between child marriage and these outcomes might be driven in part by unobservable traits or by reverse causality and therefore, the statistical associations identified in these studies should not be interpreted as the causal effects of child marriage. Relying on parental anxiety for marrying off their daughters once they reach puberty, some researchers address the endogeneity in the link between child marriage and socioeconomic outcomes through using age at menarche as an instrumental variable for age at marriage. Using this approach, a handful of studies confirm the negative effects of child marriage on women's education and health; and document the intergenerational effects on the education, health and cultural preferences of their children (Field and Ambrus, 2008; Chari et al., 2017; Sekhri and Debnath, 2014; Hicks and Hicks, 2015; Asadullah et al., 2016; Asadullah and Wahhaj, 2016). However, although this instrumental variable approach dominates the literature on the causal effects of child marriage, it is not without its critiques⁵.

With the objective of reducing the high prevalence of child marriage among Ethiopian girls, the Federal Government of Ethiopia approved in July 2000 the Revised Family Code (RFC). This law increased the legal age of marriage for women from 15 to 18 years in some regions of Ethiopia, while leaving unchanged at 18 the legal age of marriage for men. Remarkably, exposure to a legal age of marriage at 18 relative to the possibility of getting legally married at the age of 15 decreased by 20 percentage points the incidence of child marriage among the women in the sample and increased by approximately 2 years the mean age at cohabitation for these women. However, the 2011 DHS survey shows that the share of Ethiopian women cohabiting before the age of 18 remains above 20% even among those women effectively exposed to a legal age of marriage at 18. One possible cause for this is the lack of capacity of Ethiopian institutions to enforce the minimum age of marriage established in the law. The fact that the RFC banned underage marriage but not underage cohabitation could also help to explain why, even in the presence of a strong social stigma associated with unmarried cohabitation (Jones et al., 2016), setting the minimum age of marriage at 18 years has not eradicated underage cohabitation among

⁵Discussed in section 1.2.

Ethiopian women.

This study investigates the impact of the approval of the RFC in some regions of Ethiopia to pursue a twofold objective. First, I estimate the effect for women of exposure to a legal age of marriage at 18 on the probability of infant mortality of the first born child. Second, I assess the causal effect of delaying women’s age at cohabitation on the probability of infant mortality of the first born child, assessing also the mechanisms through which early cohabitation could affect infant mortality.

The novelty of the analysis presented in this study banked on three main contributions to the literature. First, this is to the best of my knowledge the first study that provides evidence on the socioeconomic consequences of increasing the minimum age of marriage for women. Second, the study focuses on the link between child marriage and infant mortality, a key development outcome that has been ignored by previous causal studies assessing the consequences of child marriage. Third, unlike previous studies relying on the use of age at menarche as an instrumental variable for child marriage, I address endogeneity between women’s age at cohabitation and socioeconomic outcomes through exploiting age discontinuities in the effective legal age of marriage faced by Ethiopian women. A regression discontinuity design (RDD) is applied exploiting the fact that those women younger than 15 years when the RFC was introduced were exposed to a legal age of marriage at 18 years, while those women that were equal or older than 15 years at the same time had the opportunity to get legally married before they turned 18 years old. This methodological approach can be also used to expand the analysis on the effects of child marriage to other outcomes of interest and to other settings where similar laws have been approved.

The results show that the probability of infant mortality of the first born is 7.9 percentage points lower among women exposed to a legal age of marriage at 18 than among women that had the opportunity to get married at the age of 15 before the approval of the law. They also suggest that even if they do not completely end child marriage, laws setting the minimum age of marriage for women at 18 years can be effective policies to reduce infant mortality. The estimates for the causal effect of early cohabitation reveal that a one-year delay in women’s age at cohabitation during teenage years decreases the probability of infant mortality of the first born by 3.8 percentage points. The results are robust to the use of different estimation techniques, bandwidths and windows for the forcing variable;

and to different placebo tests ruling out the possibility that the impact on infant mortality is driven by systematic differences between women born in different months of the year, other interventions at the national level, other legal dispositions included in the RFC or over time decreases in infant mortality.

The analysis of mechanisms indicates that the impact of raising the legal age of marriage on the infant mortality of the first born is mainly channelled through the positive effect of delaying cohabitation on the age of women at first birth. On the other hand, the analysis suggests that the effect of early cohabitation on infant mortality is not driven by any effect of the former on women’s marriage market outcomes, participation in household decisions, education or labour force participation.

The study is structured as follows. Section 1.2 reviews the literature on the socio-economic effects of child marriage. Section 1.3 discusses the incidence of child marriage in Ethiopia and presents the law that raised the legal age of marriage for women from 15 to 18 years in some regions of the country. Section 1.4 introduces the identification strategy and section 1.5 describes the data used in the main analysis. Section 1.6 presents the main results, examining their robustness to the use of alternative estimation methods, bandwidths and placebo tests. Section 1.7 investigates the channels through which early cohabitation could affect infant mortality. Finally, section 1.8 concludes the study.

1.2 Related Literature

The causes of child marriage have been extensively studied in anthropology and sociology. In the first economic study that aimed to model child marriage, Wahhaj (2015) enumerates three of the most commonly cited. First, young brides might be preferred because they are on average meeker than older ones and because they have a longer childbearing life ahead (Goody, 1990). Second, in opposition to western countries where newly married couples are expected to live without the economic support of relatives, it is very common that young couples in developing countries where the prevalence of child marriage is large are economically supported by their families, providing incentives for early marriages (Dixon, 1971). Finally, in many of these countries, the social status of the households depends strongly on the *purity* of the women of the family. In this context, families have to control the sexual behaviour of the girls of the household after sexual maturation, providing

parents incentives to marry their daughters as soon as possible after menarche (Moghadam, 2004). Consistent with the *purity* argument, Wahhaj (2015) develops a theoretical model aiming to explain child marriage in developing countries. The latter paper sets a marriage market model where women’s *purity* is noisily observed and perceived *purity* decreases with time on the marriage market, providing households strong incentives for early marriages.

Using DHS and Multiple Indicator Cluster Surveys (MICS), different studies assess the link between child marriage and socioeconomic outcomes for women and their children. A synthesis of this literature is provided in Parsons et al. (2015). The review concludes that, overall, child marriage is associated with harmful socioeconomic outcomes for women including lower levels of participation in household decision making and worse marriage market outcomes (Jensen and Thornton, 2003; Elborgh-Woytek et al., 2007; Solanke, 2015), lower labour force participation and educational attainment (Elborgh-Woytek et al., 2013; Field and Ambrus, 2008; Jensen and Thornton, 2003; Nguyen and Wodon, 2015; Wodon et al., 2016), and worse maternal health (Field and Ambrus, 2008; Campbell, 2002). The review also suggests that child marriage is associated with higher fertility, teenage pregnancy and lower age at first birth (Jensen and Thornton, 2003; Solanke, 2015). Besides, the authors of the review find that child marriage seems to be also associated with negative outcomes for the children of these women, including lower educational attainment and worse health (UNICEF, 2014; Wachs, 2008). Although not included in the review, Raj et al. (2010) and Adhikari (2003) assess empirically the statistical association between women’s age at marriage, age at first birth and mortality of young children. While the former paper provides evidence of a large and positive statistical association between infant mortality and marriage before the age of 18 in India, the latter finds that being mother before the age of 20 is strongly associated with larger levels of neonatal mortality in Nepal.

With the exception of Field and Ambrus (2008), the studies reviewed in Parsons et al. (2015) examine the correlation between child marriage and the socioeconomic outcomes of these women and their children. This could be problematic because the statistical association between child marriage and these outcomes might be driven by reverse causality or by unobservable factors correlated with both. A few studies address empirically this problem through using age at menarche as an instrumental variable for age at marriage. These studies exploit the quasi-random variation generated by a delay in the age at menarche as a source of exogenous variation for age at marriage for women. They argue that the

arrival of puberty determines the entrance in the marriage market and they prove empirically that a delay in the age at menarche increases significantly age at marriage. Using this approach, [Field and Ambrus \(2008\)](#), [Asadullah et al. \(2016\)](#) and [Hicks and Hicks \(2015\)](#) find that early marriage decreases educational attainment for women and antenatal health investments in Bangladesh, India and Kenya. Furthermore, [Chari et al. \(2017\)](#), [Sekhri and Debnath \(2014\)](#), [Asadullah et al. \(2016\)](#) and [Asadullah and Wahhaj \(2016\)](#) document for India and Bangladesh that early marriage also has intergenerational effects, leading to negative impacts on the educational attainment, cultural values and health investments received by the children of women marrying young. On the other hand, [Hicks and Hicks \(2015\)](#) do not find any effect of early marriage on labour market outcomes, beliefs and marriage market outcomes of women in Kenya; suggesting that the statistical association between child marriage and these outcomes found in studies conducting correlation analysis could be driven by reverse causality or omitted variables bias. To the best of my knowledge, this instrumental variable approach has not been used to investigate the link between child marriage and infant mortality.

The correct identification of the effect of early marriage in these studies relies on the assumption that conditional on height, socioeconomic background and location, age at menarche does not affect women's socioeconomic outcomes other than through delaying age at cohabitation. Furthermore, to be a valid instrumental variable, age at menarche should not be driven by unobservable factors that affect the outcome of interest. Although [Field and Ambrus \(2008\)](#) rule out different mechanisms through which age at menarche could affect the outcomes of interest, medical studies suggest that some childhood experiences such as stressful family environment or sexual abuse that may have long-term effects on socioeconomic outcomes seem to bring forward the age at menarche ([Karapanou and Papadimitriou, 2010](#); [Barrios et al., 2015](#)). Under the latter hypothesis, the instrumental variable used might not satisfy the exclusion restriction and the estimates in these studies should be interpreted with caution.

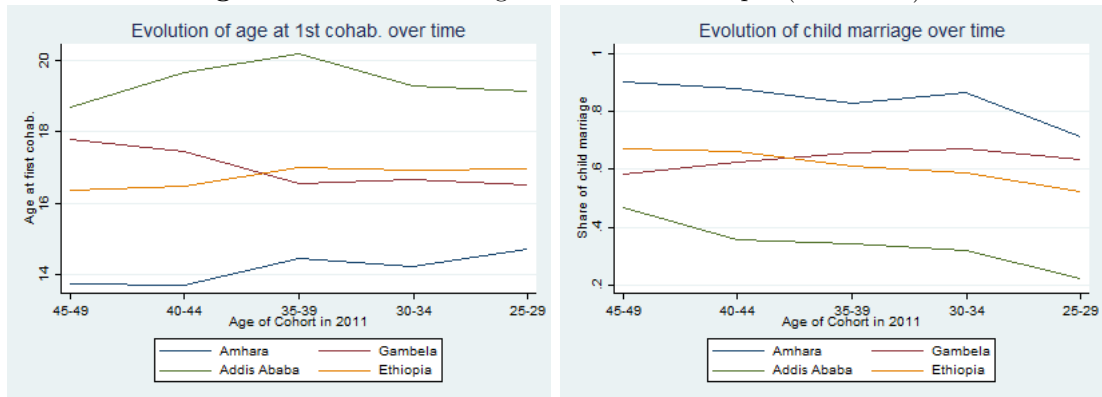
1.3 Child Marriage in Ethiopia and the Revised Family Code

The 2011 Demographic and Health Survey reveals that 41% of women aged 20-24 in Ethiopia first cohabited with a partner before the age of 18. The total number of child-

married women in the country is estimated at 1,974,000, making Ethiopia the country with the 5th largest number of child marriages in the world⁶.

Ethiopia is a federal state formed by 11 regions, which are very diverse in terms of ethnicity and religion. Although the practice of child marriage is conducted in all the country, its prevalence ranges substantially across regions. Figure 1.1 displays the evolution of the prevalence of early cohabitation and of the mean age at first cohabitation in Ethiopia. In order to illustrate the variation across regions, the figure also displays the evolution of these indicators for a selection of three Ethiopian regions, including Addis Ababa and Amhara, the regions with the largest and the lowest prevalence of child marriage in Ethiopia. For women aged 25-29 years old, the figure shows that the incidence of child marriage ranges between 70% in Amhara and 20% in Addis Ababa. Another interesting pattern that emerges from the graph is that although during the last decades the share of women cohabiting before the age of 18 has been decreasing in Ethiopia, the trend is very different across regions and while for example the incidence of child marriage has decreased sharply in Amhara or Addis Ababa, it has slightly increased in Gambela.

Figure 1.1: Child marriage over time in Ethiopia (DHS 2011)



In the last decades, increasing age at cohabitation for women has become a priority for policy makers in Ethiopia. Following the ratification of the CEDAW, which encourages governments to set and enforce laws and programmes to prevent early marriage and delay age at cohabitation, the Federal Government of Ethiopia approved the Revised Family Code (RFC) in July 2000. This law established the legal age of marriage for both men and women at 18 years.

Before the Federal Government of Ethiopia passed the RFC, the legal age of marriage for women and men was regulated by the 1960 Family Code. The latter law set a legal age

⁶Girls not Brides website.<http://www.girlsnotbrides.org/where-does-it-happen/>

for marriage of 15 years for women and of 18 years for men. Thus, while the RFC raised the legal age for marriage for women from 15 to 18 years, it left unchanged the minimum age of marriage for men at 18 years. Additionally, the RFC provided women authority to administer common marital property, abolished the right of husbands to forbid women to work outside home and facilitated the divorce procedure. The potential confounding effects of these additional dispositions are examined and dismissed in section 1.6.4. Finally, the RFC recognized the validity of marriages celebrated before the approval of the RFC that complied with the 1960 Family Code.

However, the approval of the RFC by the Federal Government of Ethiopia did not imply the immediate application of the law over the entire country. Under the Federal Constitution of Ethiopia approved after the fall of Mengistu's government, the family law is jurisdiction of the regional governments. In consequence, the approval of the RFC by the Federal Government of Ethiopia in July 2000 only implied its immediate application in the chartered cities of Addis Ababa and Dire Dawa. The application in the rest of Ethiopian regions required the approval of the regional governments. Although the enactment of the law by the Federal Government of Ethiopia paved the road, its approval by the different regional governments was not immediate (Hallward-Driemeier and Gajigo, 2015).

Figures 1.2 and 1.3 display the local linear density estimator (McCrary, 2008) of the age at first cohabitation for those women that were 12-14 years old and for those women aged 15-17 at the time of approval of the RFC in their region⁷. The figures show that while for the younger cohort of girls, exposed only to a legal age of marriage at 18 years, the most frequent age at first cohabitation is 18 years, the density function for the second cohort of women, exposed at least for some time after they turned 15 to a legal age of marriage at 15 years, reaches its peak at the age of 15 years. Furthermore, the figures suggest that while for the older cohort there seems to be a discontinuity in the density of women that first cohabited with a partner at the age of 15 and there is no discontinuity at 18 years, for the younger cohort the discontinuity at 15 years seems to be smaller and a new discontinuity emerges at the age of 18. In this line, figure 1.14 in appendix 1.B compares in the same graph the distribution of the age at cohabitation for different cohorts of women in the analytical sample. The figure shows that the distribution shift to the right

⁷These two figures are constructed using the sample of ever cohabited women in the 2011 Ethiopian DHS that were older than 18 at the time of the survey and lived in one of the five Ethiopian regions that introduced the RFC before 2008.

for the younger cohort of women, exposed to a legal age of marriage at 18 years.

Figure 1.2: Age at 1st cohabitation (cohorts 15-17 at RFC): Discontinuities at 15 and 18

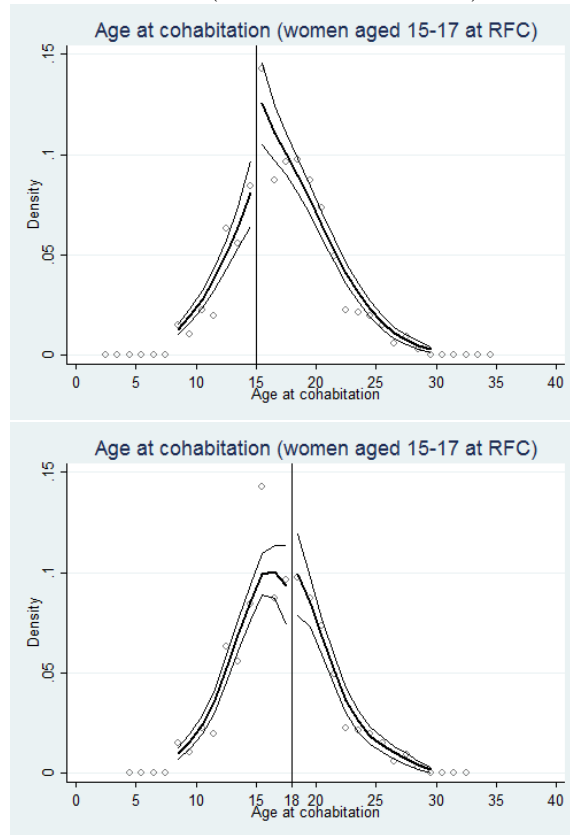
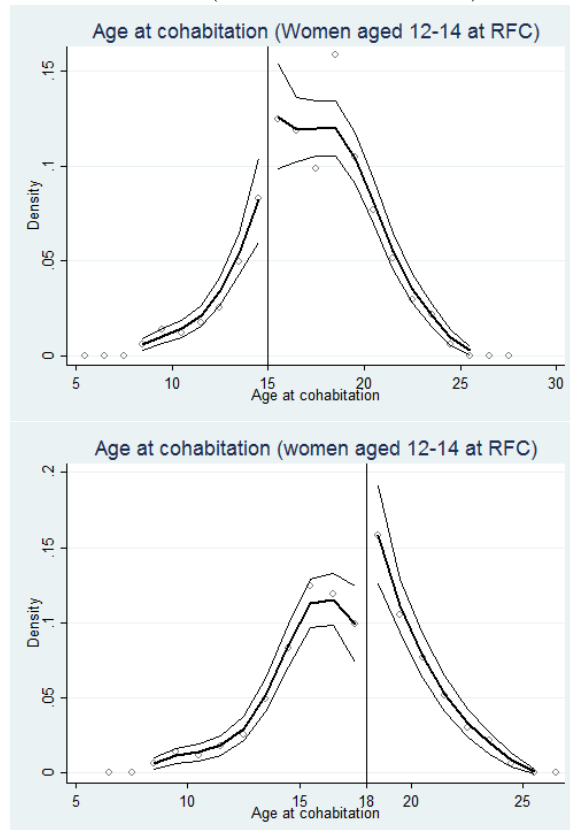


Figure 1.3: Age at 1st cohabitation (cohorts 12-14 at RFC): Discontinuities at 15 and 18



The shift in the distribution of the age at first cohabitation for these two cohorts of women suggests that the rise in the legal age of marriage increased the mean age at cohabitation for women. Indeed, the fact that the most frequent age at cohabitation in each cohort is the legal age for marriage that they are exposed to could be indicating that the minimum ages of marriage set in the 1960 and 2000 Family Codes were to some extent enforced. On the other hand, the figures confirm that the percentage of women that cohabit with a partner before reaching the minimum age of marriage is non-negligible among women from both cohorts. Different reasons can explain why the introduction of the RFC has not eradicated child marriage. First, although the RFC bans civil, religious and customary marriages before the age of 18 years, underage cohabitation is not explicitly forbidden in the law. Second, although underage marriage is not permitted in the RFC, the criminal law did not sanction it until the year 2005. Third, the institutional capacity to enforce the law is limited, particularly in rural areas where the presence of the state administration is narrow (Jones et al., 2016). Fourth, the law attributes the obligation to verify that both bride and groom are at least 18 years to the official or priest celebrating the wedding. However, the lack of birth and school registers for wide sectors of the population makes more difficult this checking procedure (Jones et al., 2016). For these reasons, and despite unmarried cohabitation is stigmatized in Ethiopia (Jones et al., 2016), we cannot expect the law to completely end child marriage.

Taken together, these patterns suggest that although it does not end with this practice, exposure to an effective legal age of marriage at 18 increases the mean age at first cohabitation for women. In section 1.6.1, I examine whether this change is sharp at the cut-off and statistically relevant to validate the estimations conducted in the study.

1.4 Identification Strategy

The RFC raised the legal age of marriage for women in Ethiopia from 15 to 18 years. This legal change generated variation in the legal age of marriage faced by women of different ages. First, those women that were younger than 15 years old when the RFC was approved were only exposed to an effective legal age for marriage at 18 years. Second, those women that were older than 18 when the RFC was approved were not directly affected by the change in the legal age for marriage. Third, those women aged between 15 and 18 years

old at the same time were exposed, at least for some time after their 15th birthday, to a legal age of marriage at 15 years. Thus, these women had the opportunity to marry legally before the age of 18. The identification strategy exploits the sharp reduction in the mean age at cohabitation with a partner for those women older than 15 years at the time of the approval of the RFC. At the extreme, the estimates rely on the change in the mean age at cohabitation for those women that were 15 years and 1 month at the time of the approval of the law; and therefore had the opportunity to get legally married at the age of 15 before the introduction of the RFC, relative to those women that were 14 years and 11 months at the same time, and could not get legally married until the age 18. If the mean age at cohabitation increases sharply for those women younger than 15 at the time of the approval of the RFC, the setting would be ideal for the implementation of a regression discontinuity design (RDD) using age of the women at the time of the approval of the RFC as the forcing variable.

The regression discontinuity framework used in this study has three particularities. First, the approval of the RFC hindered child marriage for those women younger than 18 years and not yet married when the RFC was approved but did not eradicate it among them⁸. In other words, exposure to a legal age of marriage at 18 does not fully determine in the sample whether a woman started cohabiting after the age of 18. Second, the forcing variable is defined as the age of the women at the time of the rise in the legal age of marriage measured in months, with the cut-off at the age of 15 years. I dropped from the sample used in the analysis those women that turned 15 the month in which the RFC was approved. The reason for this is that the DHS survey used in the analysis only collected information on the month and year of birth, making it impossible to determine for those women that turned 15 the same month, whether they did before or after the approval of the RFC. In this context, the quality of the birth data collected plays a crucial role enabling the use of the age at RFC as the forcing variable. At this point, it is important to mention that DHS surveys are widely used by researchers in the field of fertility and the high quality and accuracy of the information collected is showed in Pullum (2008)⁹. Third, the RFC was not applied simultaneously in every Ethiopian region. Although this variation is not exploited for identification purposes, the different timing in the application

⁸The possible causes for this are discussed in section 1.3.

⁹Furthermore, unlike in other surveys conducted in sub-Saharan Africa, the distribution of the birth data collected for the DHS in Ethiopia does not show spikes in the age of surveyed individuals ended in 0 or 5.

of the rise in the legal age for marriage across regions provides more variation in the current age of women that were approximately 15 year old when the RFC was approved in their region, making the results more generalizable.

In order to estimate (a) the effect of exposure to a legal age of marriage at 18 on women's age at first cohabitation, (b) the effect of exposure to a legal age of marriage at 18 on the probability of infant mortality of the first born child and (c) the effect of women's age at cohabitation on the probability of infant mortality of the first born child, I estimate the following three regressions:

$$Age\ at\ Cohab._i = \alpha_0 + \alpha_1(Age\ at\ RFC < 15_i) + \alpha_3 F(Age\ at\ RFC_i) + \alpha_4 X_i + \mu_i \quad (1.1)$$

$$InfantMortality_i = \delta_0 + \delta_1(Age\ at\ RFC < 15_i) + \delta_2 F(Age\ at\ RFC_i) + \delta_3 X_i + \epsilon_i \quad (1.2)$$

$$InfantMortality_i = \beta_0 + \beta_1(\widehat{Age\ at\ Cohab.}_i) + \beta_2 F(Age\ at\ RFC_i) + \beta_3 X_i + u_i \quad (1.3)$$

where $InfantMortality_i$ is a dummy variable equal to 1 if the first born child of woman i died within the first year of life, $Age\ at\ RFC < 15$ is a dummy variable that indicates whether the woman was younger than 15 when the RFC was approved in her region and therefore, was exposed to an effective legal age of marriage at 18 years. X is a vector of control variables including the region of residence, the age of the woman at the time of survey, ethnic and religion affiliation, gender of the first born and a dummy variable indicating whether the woman lives in a rural area. $F(Age\ at\ RFC)$ is a function of the age of the woman in months when the legal age for marriage was raised in her region. Finally, $\widehat{Age\ at\ Cohab.}_i$ is the predicted age at cohabitation for woman i estimated from equation 1.1.

Equation 1.1 is the first stage regression. The parameter α_1 measures the effect of exposure to a legal age of marriage at 18 on the age at first cohabitation, relative to women that had the possibility of getting legally married at 15. Equation 1.2 is the reduced form equation. The parameter δ_1 yields the effect of exposure to a legal age of marriage at 18 on the probability of infant mortality of the first born, relative to women that had the possibility of getting legally married at 15. Equation 1.3 is the second stage equation. It regresses infant mortality against the predicted age at cohabitation estimated

in equation 1.1. The parameter β_1 yields the effect of a one-year delay in women's age at cohabitation with a partner during teenage years on the probability of infant mortality of the first born.

The estimation of equations 1.1, 1.2 and 1.3 is conducted using non-parametric local polynomial regressions based on triangular kernel functions. The study follows the state of art procedure described in Calonico et al. (2014) and Calonico et al. (2016) for the selection of the optimal bandwidth and for the calculation of bias-corrected RD estimates with robust variance estimator. Standard errors are clustered at the running variable level as recommended by Lee and Lemieux (2010) for RDDs based on discrete forcing variables. As a robustness check, I also estimate equations 1.1, 1.2 and 1.3 using (a) two alternative bandwidths equal to 0.75 and 1.5 times the optimal bandwidth, (b) conventional and bias-corrected non-parametric RD estimation procedures with conventional variance estimators, and (c) parametric methods with spline polynomials of order 1 to 4 for the forcing variable and windows of 2, 3, 4 and 5 years at both sides of the cut-off.

Both when estimated using parametric and non-parametric methods, the identification of the causal effects on infant mortality of early cohabitation and of the increase in the minimum age of marriage relies on two main conditions. The first identification assumption requires that facing an effective legal age of marriage at 18 years increases the mean age at cohabitation. In other words, if the RFC did not change sharply the mean age at cohabitation for those women at the cut-off, the estimated parameter β_1 in equation 1.3 would not be efficient, potentially leading to a problem of weak instrument (Bound et al., 1995). Equally, if the reform did not change the mean age at cohabitation for the women at the cut-off, the expected coefficient of the parameter δ_1 in equation 1.2 would be 0. Although the descriptive analysis presented in section 1.3 suggests that the rise in the minimum age for marriage led to an increase in the mean age at cohabitation, the existence of a sufficiently sharp change in the mean age at cohabitation and in the incidence of child marriage at the cut-off will be tested empirically in section 1.6.1.

The second identification assumption of the RDD is that the determinants of infant mortality unaffected by the legal change should be continuously related to the forcing variable at the cut-off. Although this condition cannot be tested for every determinant of infant mortality, I examine in section 1.6.4 the existence of discontinuities at the cut-off for some of these determinants that are unlikely affected by the legal age of marriage.

If the placebo analysis shows discontinuities at the cut-off for these variables, we would need to consider the possibility that confounding factors might be driving the results. In addition to presenting the results of this placebo test, section 1.6.4 examines the robustness of the results to other identification threats and discusses the feasibility of alternative explanations for the results.

1.5 Data and Descriptive Statistics

The data used in the analysis is from the Ethiopian Demographic and Health Survey (DHS) conducted in late 2011. DHS have been implemented in more than 100 low- and middle-income countries across the world for more than three decades and they have been used in numerous studies on health and fertility in developing countries. Indeed, the vast majority of the studies and reports that explore empirically the incidence of child marriage rely on these surveys (Parsons et al., 2015). Although the questionnaires are mostly the same in DHS across the world aiming to produce comparable statistics, the exact questionnaire and the size and characteristics of the sample vary in every DHS. The 2011 Ethiopian DHS collected household, male and female level information for a sample of 16,702 households, representative at the national and regional level. The female module of the survey was applied to all women aged 15-49 living in the households sampled. This module includes questions on health, anthropometrics, demographics, fertility and women status within the household, providing information on the birth and mortality history of their children, as well as on the age at first cohabitation, which is used to measure child marriage. On the other hand, the survey provides little information on labour market outcomes and does not record the age at marriage. The questions on maternal health only target the last child of the women and data on children's health status are only recorded for those children aged 0-5 at the time of the survey.

In total, the female survey was applied to 16,515 women aged 15-49 living in the 11 Ethiopian regions. Out of these women, 8,685 live in Addis Ababa, Dire Dawa, SNNP, Tigray and Amhara, the regions that approved the RFC between 2000 and 2007 and that will be used in the analysis. The remaining 6 Ethiopian regions were excluded from the analysis for two reasons. First, four of these regions did not implement the rise in the legal age for marriage before 2008. Therefore, even if the RFC was approved in these

regions before 2011, women that were 15 when the RFC was approved would still be underage at the time of the survey. Second, the regions of Gambela and Oromia were excluded from the analysis because despite having approved the RFC before 2008, they did not seem to enforce it in any way¹⁰. Thus, the inclusion of these two regions in the analysis would decrease the magnitude and significance of the parameter of interest in the first stage equation, reducing the efficiency of the parameter that yields the effect of early cohabitation on infant mortality and potentially leading to a problem of weak instruments.

Out of the 8,685 women living in these 5 regions, I use in the analysis the sample of 5,078 women aged 18-49, that ever cohabited with a partner¹¹ and gave birth to their first child more than one year ago¹². Table 1.1 provides the descriptive statistics for the main variables used in the analysis for this sample of 5,078 women. However, it is important to remark that the regression discontinuity analysis does not use all these women to estimate the parameters of interest but only those that fall within the bandwidth used in the non-parametric analysis or the relevant window in the parametric analysis. Given that the estimates yielded by RDD are local in the sense that they are interpreted as the effects for those women that were approximately 15 when the RFC was approved in their region, the table also includes the mean of the variables for those women aged 14-15 at the time of the rise in the legal age of marriage.

The table shows that the age in 2011 for the women aged 14-15 when the RFC was approved ranges between 18 (in Tigray) and 26 (in Addis Ababa and Dire Dawa). The average number of years of education among these women is very low (less than 3), highlighting that most Ethiopian women are probably out of school by the time they start cohabiting with a partner. The participation in the labour market among the women in the sample exceeds 30% and approximately 60% of these women live in rural areas. Interestingly, 13% of women aged 14-15 when the RFC was approved in their region had separated or divorced from their first cohabiting partner.

¹⁰This pattern can be observed in figure 1.17 in appendix 1.B.

¹¹The percentage of women that gave birth without ever cohabiting with a partner only represents the 1.1% of all the women that gave birth to at least one child in these regions.

¹²Because infant mortality is defined as mortality within the first year of life, the sample is restricted to those women that gave birth to their first child more than one year ago to avoid censoring in the dependent variable.

Table 1.1: Summary statistics: Women that ever cohabited and ever bore a child in the regions included in the study.

	Aged 14-15 at RFC.					Full sample (Aged 18-49 2011)		
	N	Mean	Standard deviation	Min	Max	N	Mean	Diff (FS - 1y bw)
<i>Women characteristics</i>								
Age (2011)	308	23.38	2.28	18	26	5,078	32.88	9.50
Age at policy	308	14.58	0.49	14	15	5,078	25.36	10.78
Work (0/1)	308	0.33	0.47	0	1	5,077	0.36	0.03
Anaemia (0/1)	292	0.20	0.40	0	1	4,824	0.17	-0.03
Years schooling (compl.)	308	2.89	3.93	0	15	5,078	2.16	-0.73
Rural (0/1)	308	0.58	0.49	0	1	5,078	0.71	0.13
Muslim (0/1)	308	0.31	0.46	0	1	5,078	0.21	-0.10
Eth. Oromiya (0/1)	308	0.23	0.42	0	1	5,078	0.12	-0.11
<i>Marriage market</i>								
Age at 1st cohab	308	16.13	3.18	8	24	5,078	16.43	0.30
Child married (0/1)	308	0.66	0.47	0	1	5,078	0.65	-0.01
Currently partner (0/1)	308	0.86	0.34	0	1	5,078	0.83	-0.03
Divorced (0/1)	308	0.13	0.33	0	1	5,078	0.12	-0.01
Same partner (0/1)	308	0.73	0.45	0	1	5,076	0.62	-0.11
Empowerment index (0-2)	265	0.92	0.41	0	2	4,181	0.89	-0.03
Particip. social life decisions (0-2)	265	1.06	0.68	0	2	4,175	1.03	-0.03
Particip. health decisions (0-2)	265	1.01	0.62	0	2	4,170	0.97	-0.04
Particip. purchase decisions (0-2)	264	0.79	0.56	0	2	4,164	0.75	-0.04
Particip. husband earnings (0-2)	262	0.83	0.48	0	2	4,142	0.82	-0.01
Age difference with partner	266	6.35	4.58	-5	30	4,171	7.82	1.47
Years of schooling (partner)	307	4.12	4.59	0	16	4,999	3.53	-0.59
Wealth index (1-5)	308	3.48	1.55	1	5	5,078	3.18	-0.30
<i>Fertility outcomes</i>								
N children	308	1.85	1.01	1	7	5,078	4.22	2.37
Age at 1st birth	308	18.30	2.78	11	25	5,078	18.89	0.59
Interval cohab. first child (months)	308	27.59	26.59	0	158	5,078	31.57	3.98
<i>First born characteristics</i>								
Years since born	308	4.96	2.75	1	12	5,078	13.94	8.98
Male (0/1)	308	0.56	0.50	0	1	5,078	0.52	-0.04
Deceased before 1st year (0/1)	308	0.09	0.28	0	1	5,078	0.10	0.01
<i>Maternal and infant health: First born</i>								
N antenatal visits	107	3.39	3.49	0	12	520	3.41	0.02
Delivery at home (0/1)	164	0.48	0.50	0	1	862	0.48	0.00
N vaccines (1-9)	144	6.44	2.76	0	9	781	6.44	0.00
Postnatal check (0/1)	66	0.06	0.24	0	1	312	0.06	0.00
Months breastfeed	63	23.14	12.14	0	51	240	20.69	-2.45

Note: Descriptive statistics are provided for two different samples: (a) women aged 18-49 in the five regions of interest that ever cohabited with a partner and have given birth and (b) women aged 18-49 in the five regions of interest that were 14-15 when the RFC was approved in their region, ever cohabited with a partner and have given birth. The last column reports the difference in means between these two samples.

The mean age at cohabitation for the women in the full sample is approximately 16.4 years, and more than 60% of these women first cohabited with their partner before the age of 18 years. If we focus on the sample of individuals aged 14-15 when the RFC was introduced in their region, the incidence of child marriage falls from 70% among those women aged 15 at the time of the RFC to 60% among those women aged 14 and therefore, exposed to a legal age of marriage at 18 years. Similarly, the mean age at first cohabitation

increases from 15.8 to 16.6 for the same groups of women¹³. The differences in terms of age at cohabitation and prevalence of child marriage between women aged 14 and women aged 15 at the time of the RFC are statistically significant at the 5% and 10% significance levels.

In the full sample, the mean age of women at first birth is 18.9 years and the average interval between cohabitation and first birth is 31.6 months. The mean age at birth and the mean interval between cohabitation and first birth are smaller (18.3 and 27.6) among those women aged 14-15 when the RFC was introduced in their region. The infant mortality rate of the first born among the women in the full sample is 10%. If we focus on the women aged 14-15 when the RFC was introduced in their region, the probability of infant mortality of the first born falls from 12% among the women aged 15 when the RFC was introduced to 5% among the women aged 14 at the same time and therefore, exposed to a legal age of marriage at 18 years. This difference is statistically significant at the 5% level of significance.

Figure 1.4: Age at cohabitation, age at birth and infant mortality of the first born (LOWESS regressions)

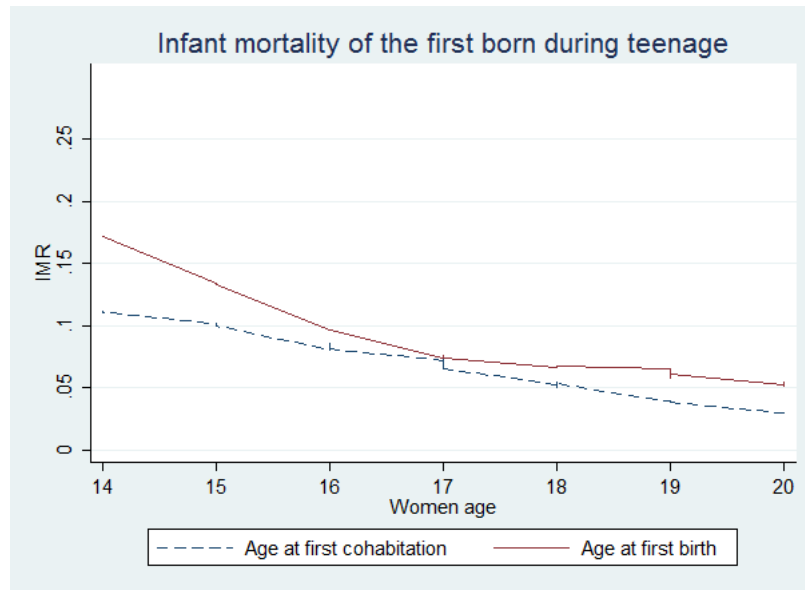


Figure 1.4 displays the statistical association between women's age at cohabitation, age

¹³The large incidence of child marriage and the low age at cohabitation among the youngest women in the sample, which represent an important part of the women within one year from the cut-off, is mechanically driven by the way in which the sample is selected. Because the sample only includes women aged 18-49 that have given birth to their first child more than one year ago and have ever cohabited, it is very likely that the vast majority of the youngest women in the sample (e.g. aged 18 or 19) cohabited with their partner before the age of 18. The evolution of the prevalence of child marriage across age cohorts in Ethiopia could be better observed in figure 1.12 in appendix 1.B, which is constructed using the women aged 18-49 in the DHS data regardless of whether they ever cohabited or gave birth.

at first birth and infant mortality of the first born during teenage years. The statistical relation between the variables is estimated using LOWESS regressions for the sample of women used in the analysis. Although these relations should not be interpreted as causal, the figure shows a strong correlation between age at first birth and infant mortality during puberty. The graph suggests that delaying age at first birth from 15 to 17 is associated with a decrease in the incidence of infant mortality of the first born from approximately 13.5% to 7.5%. On the other hand, rises in the age at first birth after the age of 17 are associated with smaller reductions in the probability of infant mortality of the first born. Although the slope of the estimated function that displays the statistical association between age at cohabitation and infant mortality of the first born during early adolescence is less pronounced, the negative statistical association between these two variables is also evident in the graph.

1.6 Results

1.6.1 The Effect of the RFC on the Age at Cohabitation

The first condition for the validity of the identification strategy outlined in section 1.4 for the estimation of (a) the effect of exposure to a legal age of marriage at 18 on the probability of infant mortality of the first born and of (b) the effect of the age at cohabitation on infant mortality of the first born, is the existence of a discrete change in the mean age at first cohabitation at the cut-off. The size and statistical significance of this discontinuity is yielded by the parameter α_1 in the first stage equation. Columns 1, 3 and 5 of table 1.2 report the estimates for this parameter using non-parametric techniques with different estimation procedures and bandwidths. The results of the preferred estimation are reported in column 3 and show that exposure to a legal age of marriage at 18 relative to the possibility of getting legally married at 15 increases women's age at cohabitation by approximately 2 years. The coefficients of the variable when the alternative bandwidths and non-parametric estimation procedures are used are also large and positive. Overall, the coefficients measuring the effect of exposure to a legal age of marriage at 18 across the different non-parametric estimations are statistically significant at the 1% and satisfy the *relevance* condition ($F > 10$) required for the estimation of the second stage equation.

The results reported in columns 1, 3, 5 and 7 of table 1.3 show that the rise in the

age at first cohabitation for women younger than 15 when the RFC was implemented in their region remains large (0.8-1.8 years) and statistically significant when equation 1.1 is estimated using parametric methods with several windows and spline polynomials for the forcing variable. The sharp change in the mean age at cohabitation at the cut-off is also evident in figure 1.5. Consistently, the results reported in column 1 of table 1.4 show that the rise in the mean age at cohabitation is accompanied by a decrease in 20 percentage points in the incidence of child marriage at the cut-off. The results of the first stage equation are in line with the main conclusions of the descriptive analysis conducted in section 1.3 showing how the distribution of the age at cohabitation changes across cohorts of women exposed to a different legal age of marriage.

Figure 1.5: Main analysis: Age at first cohabitation at the cut-off

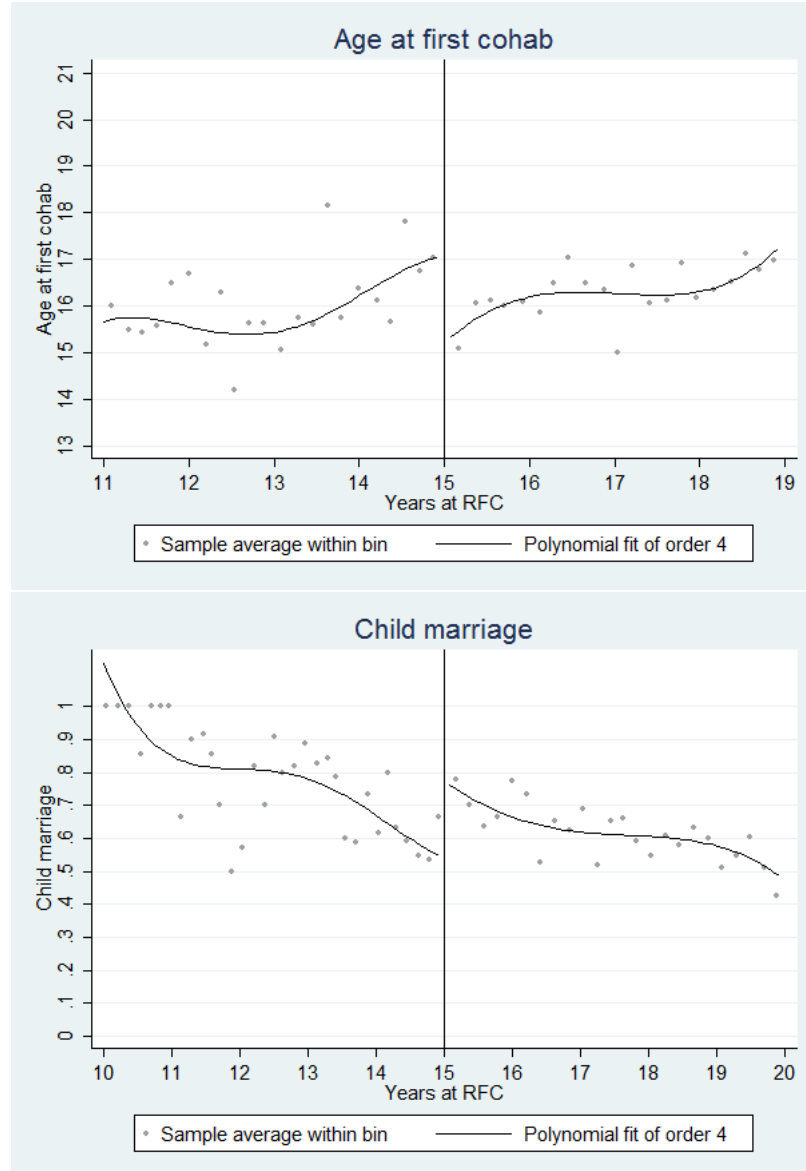


Table 1.2: Non-parametric methods: RFC, age at first cohabitation and infant mortality.

	Conventional		Bias-corrected		Robust	
	(1) FS Age at 1st cohab	(2) SS/RF Infant Mortality	(3) FS Age at 1st cohab	(4) SS/RF Infant Mortality	(5) FS Age at 1st cohab	(6) SS/RF Infant Mortality
<i>Bandwith A: Calonico et al. (2016)</i>						
Age<15 at RFC	1.774*** (0.000)	-0.073** (0.034)	2.055*** (0.000)	-0.079** (0.021)	2.055*** (0.000)	-0.079* (0.051)
Age at 1st cohab.		-0.041** (0.031)		-0.038** (0.044)		-0.038* (0.095)
N		5078		5078		5078
N effect. obs.		581		990		990
Bandwidth		24.0		40.3		40.3
<i>Bandwith B: $1.5 \times C C T$</i>						
Age<15 at RFC	1.380*** (0.000)	-0.073*** (0.007)	1.584*** (0.000)	-0.082*** (0.002)	1.584*** (0.000)	-0.082*** (0.008)
Age at 1st cohab.		-0.053*** (0.004)		-0.052*** (0.005)		-0.052** (0.016)
N		5078		5078		5078
N effect. obs.		874		1451		1451
Bandwidth		36.1		60.5		60.5
<i>Bandwith C: $0.75 \times C C T$</i>						
Age<15 at RFC	1.869*** (0.000)	-0.078* (0.074)	2.081*** (0.000)	-0.082* (0.059)	2.081*** (0.000)	-0.082 (0.122)
Age at 1st cohab.		-0.042* (0.068)		-0.039* (0.086)		-0.039 (0.172)
N		5078		5078		5078
N effect. obs.		453		733		733
Bandwidth		18.0		30.2		30.2

Note: Each coefficient provided in the table is estimated using a separate regression. The table reports the estimates of interest for the first stage (FS), reduced form (RF) and second stage (SS) equations using different bandwidths and the three procedures described in Calonico et al. (2016): conventional variance estimator, bias-corrected variance estimator and robust variance estimator. The coefficients for the variable *Age<15 at RFC* measure the effect of the RFC on the age at first cohabitation (first stage) in columns 1, 3 and 5; and the effect of the law on the prevalence of infant mortality (reduced form) in columns 2, 4 and 6. The coefficients for the variable *Age at 1st cohab* measure the effect of delaying one year the age at cohabitation during teenage years on the prevalence of infant mortality (second stage equation). The results provided for Bandwidth A are estimated using the optimal bandwidth calculated following Calonico et al. (2016). The results provided for Bandwidths B and C are estimated using $1.5 \times$ the optimal bandwidth and $0.75 \times$ the optimal bandwidth. The specifications include the following control variables: dummies for the regions of residence, the age of women at survey, ethnic and religion affiliation, gender of the first born, a rural/urban dummy variable and a non-parametric function for the age of the women at RFC. The sample size and the bandwidths used in the RF, FS and SS regressions are the same within each estimation procedure and bandwidth used. Standard errors are clustered at the forcing variable. P-values are in parentheses. ***p<0.01; **p<0.05; *p<0.1.

The evolution of the prevalence of child marriage and mean age at cohabitation across age cohorts observed in figure 1.5 deserves two additional comments. First, the large discontinuity at the cut-off and the polynomial behaviour at the right of the cut-off are consistent with the hypothesis that through pushing women slightly over 15 when the RFC was introduced to get married before the approval of the RFC, the introduction of the RFC could have reduced the mean age at cohabitation for the cohorts of women aged

just above 15 at the time of the RFC. For example, a woman slightly older than 15 years old when the RFC was approved had the possibility of getting legally married as soon as she turned 15, but if she waited some months and the RFC is approved, she would not be able to get legally married until the age of 18. This fact might push women aged 15 at the time of the RFC that were planning to get married over the next 2-3 years to marry as soon as they turned 15. This fact has an important implication for the interpretation of the results: the estimates of interest in the first stage and reduced form equations measure the effect of exposure to a legal age of marriage at 18 relative to the possibility of getting married at 15, rather than to exposure to a legal age of marriage at 15.

Table 1.3: Parametric methods using different time windows: RFC, age at first cohabitation and infant mortality.

	2 years window (N= 571)		3 years window (N= 849)		4 years window (N= 1,164)		5 years window (N= 1,432)	
	(1) FS Age at 1st cohab	(2) SS/RF Infant mortality	(3) FS Age at 1st cohab	(4) SS/RF Infant mortality	(5) FS Age at 1st cohab	(6) SS/RF Infant mortality	(7) FS Age at 1st cohab	(8) SS/RF Infant mortality
<i>Spline Pol. order 1</i>								
Age<15 at RFC	1.707*** (0.000)	-0.076** (0.022)	1.057*** (0.001)	-0.076** (0.022)	0.919*** (0.001)	-0.053** (0.041)	0.849*** (0.002)	-0.047** (0.040)
Age at 1st cohab		-0.045** (0.012)		-0.045** (0.012)		-0.058** (0.041)		-0.055** (0.041)
<i>Spline Pol. order 2</i>								
Age<15 at RFC	1.552*** (0.000)	-0.080* (0.054)	1.565*** (0.000)	-0.070** (0.035)	1.335*** (0.000)	-0.068** (0.029)	1.159*** (0.000)	-0.058* (0.051)
Age at 1st cohab		-0.052** (0.040)		-0.045** (0.025)		-0.051** (0.021)		-0.050** (0.039)
<i>Spline Pol. order 3</i>								
Age<15 at RFC	1.552*** (0.000)	-0.080* (0.054)	1.552*** (0.000)	-0.097** (0.018)	1.321*** (0.000)	-0.069** (0.036)	1.106*** (0.001)	-0.058* (0.064)
Age at 1st cohab		-0.052** (0.040)		-0.057*** (0.010)		-0.052** (0.032)		-0.052* (0.055)
<i>Spline Pol. order 4</i>								
Age<15 at RFC	1.552*** (0.000)	-0.080* (0.054)	1.552*** (0.000)	-0.097** (0.018)	1.782*** (0.003)	-0.100* (0.061)	1.360*** (0.002)	-0.082** (0.032)
Age at 1st cohab		-0.052** (0.040)		-0.057*** (0.010)		-0.056** (0.043)		-0.060** (0.030)

Note: Each coefficient provided in the table is estimated using a separate regression. The table reports the estimates of interest for the first stage (FS), reduced form (RF) and second stage (SS) equations using the sample of women aged within different age windows around the cut-off and spline polynomials of order 1 to 4 for the forcing variable. The coefficients for the variable *Age<15 at RFC* measure the effect of the RFC on the age at first cohabitation (first stage) in columns 1, 3, 5 and 7; and the effect of the law on the prevalence of infant mortality (reduced form) in columns 2, 4, 6 and 8. The coefficients for the variable *Age at 1st cohab* measure the effect of delaying one year the age at cohabitation during teenage years on the prevalence of infant mortality (second stage equation). The specifications include the following control variables: dummies for the regions of residence, the age of women at survey, ethnic and religion affiliation, gender of the first born, a rural/urban dummy variable and a spline polynomial function of order 1 to 4 for the age of the women at RFC. The order of the polynomial function for the forcing variable used in every specification is reported above each estimation set. Standard errors are clustered at the forcing variable. P-values are in parentheses. ***p<0.01; **p<0.05; *p<0.1.

Second, beyond the large discontinuity at the cut-off, the figure suggests that the prevalence of child marriage is larger among the younger women in the sample, exposed to

a legal age of marriage at 18. Although we cannot expect child marriage to be eradicated among women exposed to a legal age of marriage at 18, the fact that the largest incidence of child marriage in the sample is found for the youngest cohorts of women is apparently puzzling. To reconcile this paradox, one should take into account that the sample used in the analysis only includes women that ever cohabited with a partner and have given birth to their first child at least one year before the survey. In this scenario, the prevalence of child marriage among the youngest cohorts of women in the sample, barely aged 18 at the time of the survey, is expected to be very close to 100%. The *true* evolution of the prevalence of child marriage across age cohorts in Ethiopia is presented in figure 1.12 in appendix 1.B. Using the full sample of women aged 18-49 included in the DHS data regardless of whether they ever cohabited with a partner or have given birth, the figure shows that the prevalence of child marriage is lower among younger cohorts of women, and changed dramatically at the cut off.

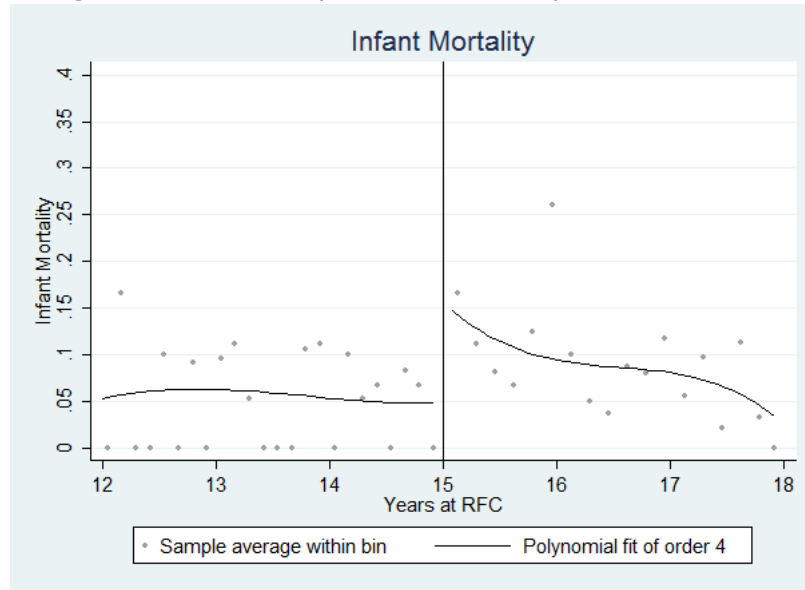
1.6.2 The Effect of the RFC on Infant Mortality

The next step is determining whether the 20 percentage points drop in the incidence of child marriage and the 2 years increase in the mean age at cohabitation at the cut-off for women exposed to a legal age of marriage at 18 affected the probability of infant mortality of the first born. This effect is yielded by the parameter δ_1 in equation 1.2. The estimates for this parameter using non-parametric estimations with different estimation procedures and bandwidths are reported in columns 2, 4 and 6 of table 1.2. The coefficients suggest that exposure to a minimum age of marriage at 18 years relative to the possibility of getting legally married at 15, decreases significantly the probability of infant mortality of the first born by 7.3-8.2 percentage points, depending on the bandwidth and estimation procedure selected. The preferred estimate is reported in column 6 and yields an effect of 7.9 percentage points, statistically significant at the 10%. Columns 2, 4, 6 and 8 of table 1.3 provide the estimates for the parameter δ_1 using parametric techniques with different windows and spline polynomials of order 1 to 4 for the forcing variable. Overall, the estimates presented in these two tables confirm the robustness of the results to the use of parametric and non-parametric methods with different bandwidths, windows and polynomial functions for the forcing variable. The discontinuity in the infant mortality rate of the first born for those women older than 15 at the time of the approval of the

RFC is graphically displayed in figure 1.6.

The magnitude of the effect of exposure to a legal age of marriage at 18 on the infant mortality of the first born at the cut-off seems large, particularly when compared with the mean incidence of infant mortality of the first born among the women in the sample (10%). However, the interpretation of this coefficient requires a few considerations. First, women just above the cut-off had the opportunity of getting legally married as soon as they turned 15, but if they wait a few months and the RFC is approved, they would face a legal age of marriage at 18. The main implication for this, discussed in the next to last paragraph in section 1.6.1, is that the parameter of interest in the reduced form equation measures the effect on infant mortality of the first born of exposure to a legal age of marriage at 18 relative to the possibility of getting legally married at the age of 15, rather than to exposure to a legal age of marriage at 15 years.

Figure 1.6: Main analysis: Infant mortality rate at the cut-off



Second, the RDD estimates of the effects on infant mortality of the first born presented in this section are larger but aligned with those obtained in simple correlation analysis. For example, the results of the LOWESS analysis displayed in figure 1.4 reveal that delaying the age of women at first cohabitation from 15 to 17, which is the approximate change in mean age at cohabitation at the cut-off, would be associated with a decrease of approximately 3.5-4 percentage points in the incidence of infant mortality of the first born.

Third, the estimates of the parameter δ_1 discussed in this section should be interpreted as local treatment effects: they measure the effect on infant mortality of exposure to a

minimum age of marriage at 18 for those women in the sample that were approximately 15 years old when the RFC was approved in their region. These are women that when they were surveyed in 2011, had an age ranging between 25 (Addis Ababa and Dire Dawa) and 18 years (Tigray), have already cohabited with a partner and have given birth to their first child more than one year before the survey.

1.6.3 The Effect of Women's Age at Cohabitation on Infant Mortality

The causal effect of women's age at cohabitation during teenage years on the infant mortality of the first born is yielded by the parameter β_1 in the second stage equation. The results for the non-parametric estimations are reported in columns 2, 4 and 6 of table 1.2 and reveal that a one-year delay in the age at first cohabitation decreases the probability of infant mortality of the first born by 3.8-5.2 percentage points, depending on the bandwidth and estimation procedure used. The preferred estimation, reported in column 3, indicates that a one-year delay in the age of women at cohabitation decreases the probability of infant mortality of the first born by 3.8 percentage points. The effect is statistically significant at the 90% confidence level. The results for the effect of women's age at cohabitation reported in columns 2, 4, 6 and 8 of table 1.3 confirm that the findings of the non-parametric analysis are robust to the use of parametric techniques with different windows and spline polynomials of order 1 to 4 for the forcing variable.

When interpreting the coefficients of the second stage equation, it is important to consider that the estimated effect of early cohabitation on the infant mortality of the first born is a local average treatment effect. More specifically, the parameter of interest in the regression measures the effect of a one-year delay in the age at cohabitation during teenage years for those women in the sample that were approximately 15 years old when the RFC was approved in their region and delayed cohabitation because they were exposed to a legal age of marriage at 18 years.

1.6.4 Robustness Checks

Using parametric and non-parametric methods, the previous section shows that exposure to a minimum age of marriage at 18 relative to the possibility of getting legally married at 15 increases significantly age at cohabitation and decreases the incidence of child marriage and infant mortality of the first born. In this section, I discuss and explore alternative

explanations for the results.

Firstly, one important threat to the attribution of the effects identified on infant mortality to the rise in the legal age for marriage would be that the RFC not only raised the legal age for marriage but also set some additional provisions aiming to change the balance of power within the household through facilitating the procedure of divorce, abolishing the right of husbands to forbid women to work and providing women the right to administer the common marital property. These legal changes may have improved women economic status, empowerment and participation in household decisions ultimately affecting infant mortality.

Table 1.4: RFC and different outcomes.

	(1) Child marriage	(2) Age at 1st cohabit.	(3) Paid work	(4) Empowerment index	(5) Decision relative visits	(6) Decision HH purchases	(7) Decision health	(8) Decision husband earn	(9) Divorced (Only ever cohab)
Age<15 at RFC	-0.200*** (0.000)	2.094*** (0.000)	0.098 (0.211)	-0.054 (0.395)	-0.318** (0.014)	0.132 (0.247)	-0.059 (0.643)	-0.063 (0.275)	0.018 (0.686)
N	5078	5078	5077	4181	4175	4164	4170	4142	5078
N effect. obs.	958	812	1176	1035	801	771	898	893	958
Bandwidth	39.2	33.1	48.1	49.3	38.0	37.4	42.2	42.9	39.5

Note: Each coefficient provided in the table is estimated using a separate regression. The table reports the estimates of interest for the reduced form (RF) equation using the optimal bandwidth and the robust variance estimator described in [Calonico et al. \(2016\)](#). The coefficients for the variable *Age<15 at RFC* measure the effect of the RFC on the outcome variable analyzed. The regressions conducted include as control variables a set of dummies for the regions of residence, the age of women at survey, ethnic and religion affiliation, gender of the first born, a rural/urban dummy variable and a non-parametric function for the age of the women at RFC. Standard errors are clustered at the forcing variable. P-values are in parentheses. ***p<0.01; **p<0.05; *p<0.1.

Although this seems a plausible possibility, the point here is that all of these norms were applied retrospectively regardless of whether women were already married or not, and therefore, they should not affect differently women aged just below and above 15 years when the RFC was approved. Nonetheless, I examine empirically whether the law affected differently labor force participation, divorce rates and participation in household decisions of women at both sides of the cut-off. To measure woman participation in household decisions, I use a set of questions in the DHS survey that provide information on who decides on relative visits, household purchases, health expenditure and the administration of the money earned by the husband. Each of these variables take the value of 0 if the woman does not participate in the decision, 1 if the woman participates in the decision and 2 if the decision is taken alone by the woman. Then, I construct a self-reported empowerment index for each woman as an average score in these questions. Using each

of these self-reported empowerment measures, labor force participation and divorce rates as dependent variables, I estimate equation 1.2. The results reported in table 1.4 suggest that, overall, the new dispositions included in the RFC aiming to change the balance of power within the household did not affect differently labor force participation, divorce rates and participation in household decisions of women at both sides of the cut-off. In consequence, the evidence confirms that the effect on infant mortality identified in the study is not driven by these additional norms aiming to improve women's bargaining power within the household.

Secondly, I examine the existence of discontinuities in variables that are plausibly not affected by the reform including the ethnicity and religion of the women and the gender of their first born. This is an indirect empirical test for the second identification assumption discussed in section 1.4: the determinants of infant mortality should be continuously related to the forcing variable at the cut-off. In order to test this hypothesis, I estimate equations 1.1, 1.2 and 1.3 using the bias-corrected RD estimates with robust variance estimator and an optimal bandwidth calculated following Calonico et al. (2014) and whether the first born is male, whether the mother is Muslim or from Oromo ethnic group¹⁴ as outcome variables. The results of these estimations are provided in columns 1-6 of table 1.5. The coefficients are small and largely insignificant, confirming that there is not any discontinuity in the value of these placebo variables at the cut-off. The absence of discontinuities at the cut-off for these placebo variables is also evident in figure 1.7.

Thirdly, I examine whether the difference in infant mortality rates of the first born among women at both sides of the cut-off could be driven by systematic differences between women born in different months of the year rather than by exposure to a different minimum age of marriage. To assess this possibility, I re-estimate equations 1.1, 1.2 and 1.3 setting a placebo cut-off for women older than 19, rather than 15, at the time of the approval of the RFC¹⁵. This exercise is equivalent to placing the cut-off as if the law was applied exactly four years before the true date of approval (e.g. 4th of July of 1996 for Addis Ababa and

¹⁴Oromo is the most prevalent ethnic group in Ethiopia.

¹⁵The set of the placebo cut-off at 19 years is driven by the convenience of setting the false cut-off at a value of the forcing variable that left out of the estimation the observations around the real cut-off. Nonetheless, I have also conducted the analysis setting the placebo cut-off for women aged 16, 17 and 18 at the time of the approval of the RFC. This exercise is equivalent to placing the cut-off as if the law was introduced exactly 1, 2 and 3 years before the true date of approval. Consistently, the results of these placebo analyses, not reported in the study, show no statistically significant discontinuities in infant mortality and age at cohabitation at the false cut-offs when the bandwidth drop from the estimation the observations in the real cut-off

Dire Dawa, etc). If the results of the study on infant mortality are driven by systematic differences between women born in different months, we would expect a discontinuity in the infant mortality rate of the first born among women born in different months every year. The results of this placebo test are reported in columns 7 and 8 of table 1.5 and they reveal no discontinuities for the mean age at cohabitation or the infant mortality rate at the false cut-off, confirming that the main conclusions of the study are not driven by systematic differences between women born in different months of the year.

Table 1.5: Robustness checks Infant Mortality: Placebo analyses.

	Placebo outcomes Ethnicity		Gender		Religion	
	(1)	(2)	(3)	(4)	(5)	(6)
	FS Age at 1st cohab	SS/RF Ethnic. Oromo	FS Age at 1st cohab	SS/RF Male	FS Age at 1st cohab	SS/RF Muslim
Age<15 at RFC	1.984*** (0.000)	-0.015 (0.762)	2.068*** (0.000)	-0.039 (0.727)	1.999*** (0.000)	0.037 (0.598)
Age at 1st cohab.		-0.006 (0.834)		-0.018 (0.755)		0.017 (0.648)
N		5078		5078		5078
N effect. obs.		874		874		874
Bandwidth		36.6		36.9		36.2
	Placebo: RFC 48 months before		Placebo: Other Ethiopian regions		Control for year first born	
	(7)	(8)	(9)	(10)	(11)	(12)
	FS Age at 1st cohab	SS/RF Infant Mortality	FS Age at 1st cohab	SS/RF Infant Mortality	FS Age at 1st cohab	Infant Mortality
Age<15 at RFC	0.167 (0.666)	-0.001 (0.961)	-0.469 (0.213)	-0.007 (0.789)	1.107*** (0.000)	-0.074** (0.023)
Age at 1st cohab.		-0.007 (0.963)		0.015 (0.810)		-0.066* (0.053)
N		5078		2398		5078
N effect. obs.		2733		1458		1286
Bandwidth		89.5		106.9		53.1

Note: Each coefficient provided in the table is estimated using a separate regression. The table reports the estimates of interest for the first stage (FS), reduced form (RF) and second stage (SS) equations using the optimal bandwidth and the robust variance estimator described in Calonico et al. (2016). The coefficients for the variable *Age<15 at RFC* measure the effect of the RFC on the age at first cohabitation (FS) and on the outcome variable analyzed (RF). The coefficients for the variable *Age at 1st cohab* measure the effect of delaying one year the age at cohabitation during teenage years on the outcome variable analyzed (SS). The sample size and the bandwidths used in the RF, FS and SS regressions are common within every outcome analyzed. Columns 1-6 report the results for placebo variables, arguably unaffected by the RFC. Columns 7 and 8 report the results of a placebo test with a false cut-off set 4 years before the approval of the RFC. Columns 9 and 10 report the results of a placebo test using the regions that did not approve the RFC by the time of the survey: Affar, Harari and Gumuz. The regressions conducted for the estimation of the coefficients reported in columns 1 to 10 include as control variables a set of dummies for the regions of residence, the age of women at survey, ethnic and religion affiliation, gender of the first born, a rural/urban dummy variable and a non-parametric function for the age of the women at RFC. Ethnic, gender and religion are not included as control variables in the regressions where these variables are the outcome variables. Columns 11 and 12 report the results for the infant mortality including as an additional control variable a time trend for the year of birth of the infant, aiming to account for over time changes in the incidence of infant mortality. Standard errors are clustered at the forcing variable. P-values are in parentheses. ***p<0.01; **p<0.05; *p<0.1.

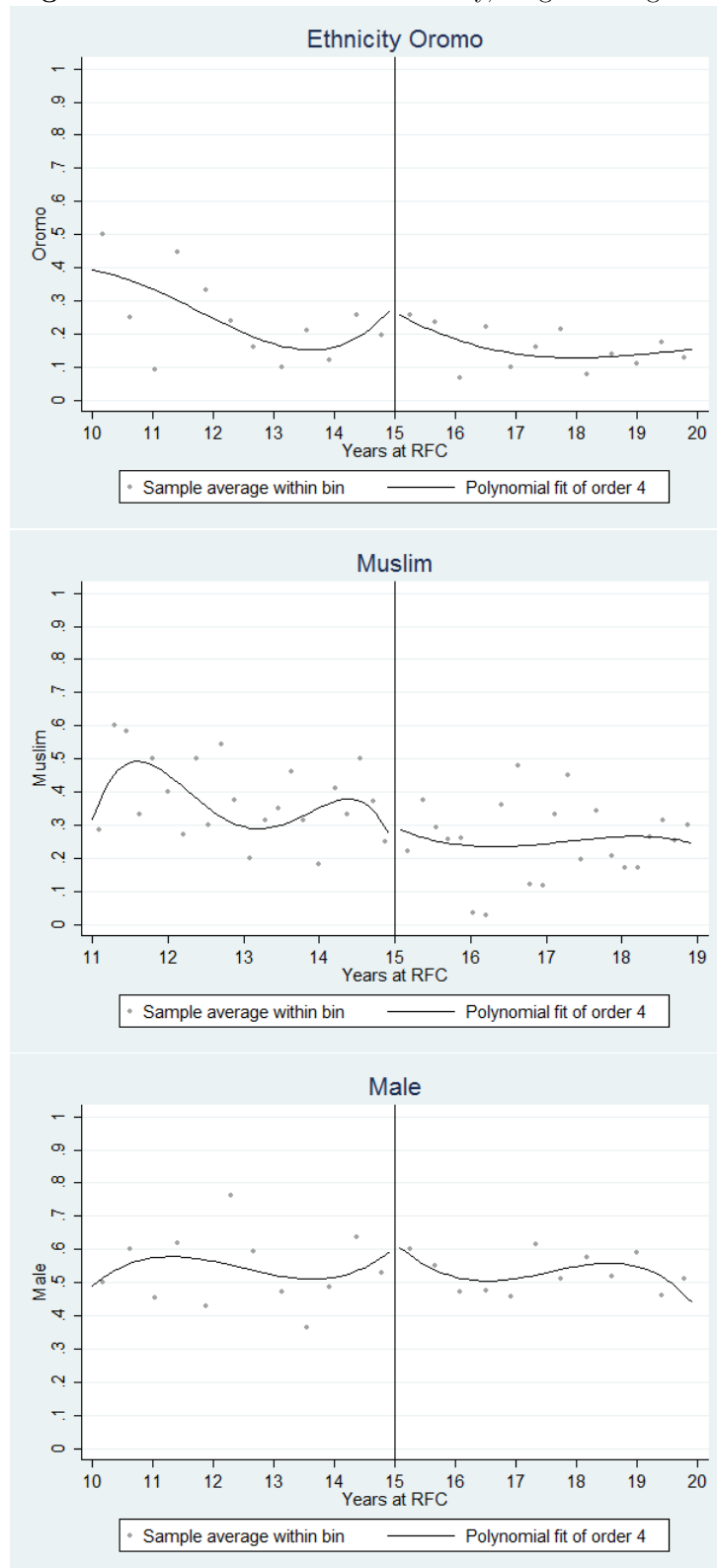
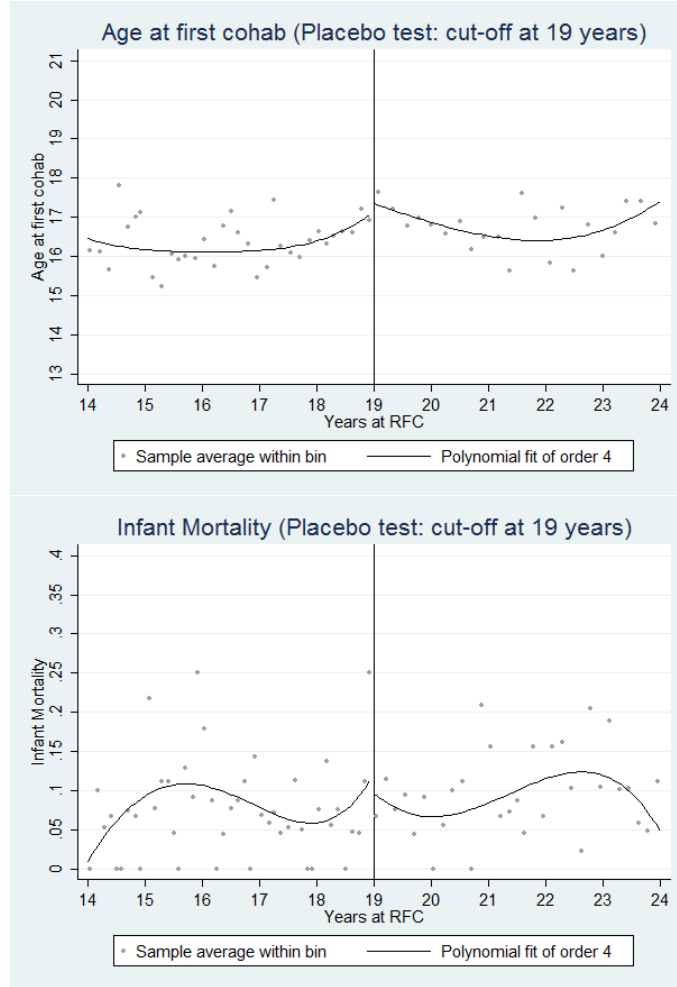
Figure 1.7: Placebo variables: ethnicity, religion and gender

Figure 1.8: Placebo test: Cut-off at 19 years

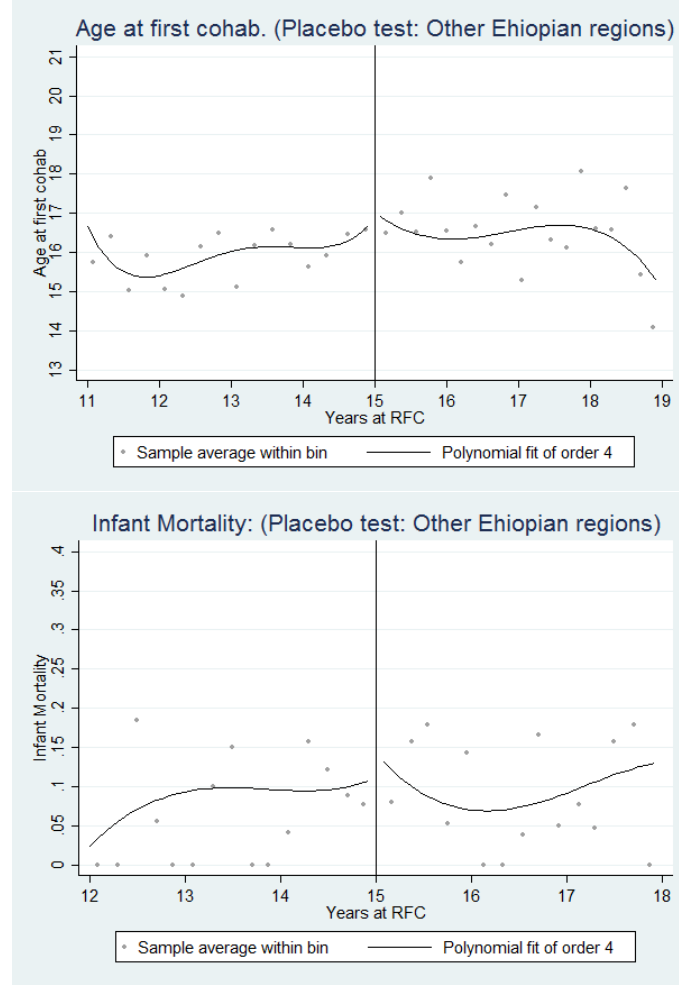
Fourthly, it is also possible that the parameter could be capturing an *odd* time trend or the effect of national level policies affecting differently women at both sides of the cut-off. In order to test this hypothesis, I re-estimate the results using the Ethiopian regions of Afar, Harari and Gumuz, setting falsely the approval date of the RFC in these placebo regions in July 2000¹⁶. I restrict the analysis to these three regions because none of them passed the RFC before 2008¹⁷ and therefore, the women used in the sample (older than 18 years at the time of the survey) were unaffected by the rise in the legal age at marriage in these regions. The results of this placebo test are reported in columns 9 and 10 of table 1.5 and they show that there is not any significant discontinuity in the mean age at cohabitation or in the infant mortality at the cut-off in those regions that have not approved the RFC, ruling out the possibility that the effect is driven by a national level

¹⁶The 4th of July of 2000 the Federal Government of Ethiopia approved the RFC and the law started to be applied in Addis Ababa and Dire Dawa.

¹⁷I exclude the region of Somali because it is unclear their exact date of introduction (Hallward-Driemeier and Gajigo, 2015)

policy affecting differently women at both sides of the cut-off.

Figure 1.9: Placebo test: Discontinuity in other Ethiopian regions



Fifthly, since the law plausibly affected the age at first birth of women at the cut-off, it is very likely that those women aged 14 years and 11 months at the time of the approval of the RFC ended up having their first child significantly later than those aged 15 years and 1 month. In this context, the parameters δ_1 and β_1 could be only capturing over-time decreases in infant mortality unrelated with the age at first cohabitation. I investigate this possibility through re-estimating equations 1.1, 1.2 and 1.3 including the year of birth of the first born as a control variable. The inclusion of this variable is aiming to capture time trends in infant mortality¹⁸. The estimates provided in columns 10 and 11 show that the direction and significance of the parameters estimated in sections 1.6.1, 1.6.2 and 1.6.3 do not vary when the year of birth of the first child is included in the regression. The latter

¹⁸The year of birth of the first born is not included as control variable in the main results reported in section 1.6 because the year of birth could be to some extent a fertility decision of the mother and therefore, plausibly affected by the age at cohabitation. Thus, including it in the main regression could lead to a *bad* control problem (Angrist and Pischke, 2008).

suggests that the reduction in the rate of infant mortality of the first born at the cut-off is not entirely driven by over time decreases in infant mortality.

Sixthly, the main results of the study are also robust to restricting the analysis to the subsample of women that by the time of the survey were still living with their first partner. The results of this robustness check are reported in columns 3 and 4 of table 1.7 in appendix 1.C.

Another potential threat to the interpretation of the estimates could be the possible existence of selective migration. In other words, those women below the age of 15 when the legal age of marriage was raised in their region that were particularly interested in early cohabitation could have migrated to regions that did not raise the legal age of marriage. If the share of women migrating for this reason is substantial, the results might be biased by selective attrition at one side of the cut-off. Although the lack of information on women's region of origin hindered the assessment of this hypothesis, the fact that the incidence of child marriage is above 10% in every Ethiopian region even among those women exposed to a legal age of marriage at 18 years could be indicating that those women (or women's families) particularly interested in cohabiting before the age of 18 may not need to migrate to another region to do so. Furthermore, the lack of discontinuity in the density of the forcing variable at the cut-off evident in figure 1.11 in appendix 1.A suggests that the migration of women slightly younger than 15 when the RFC was approved was not a widespread phenomenon.

I also explore the robustness of the results to the use of neonatal mortality as a dependent variable rather than infant mortality. For this, I re-estimate equations 1.1, 1.2 and 1.3 focusing on mortality of the first born within the first month of life, rather than within the first year. The estimates obtained for neonatal mortality of the first born have the same sign and statistical significance although the magnitudes are even larger than those obtained for infant mortality of the first born¹⁹. Therefore, the evidence confirms that the results of the study are robust to the definition of the dependent variable and suggests that most of the effect of child marriage on the mortality of the first born occurs during the first month of life of the newborn.

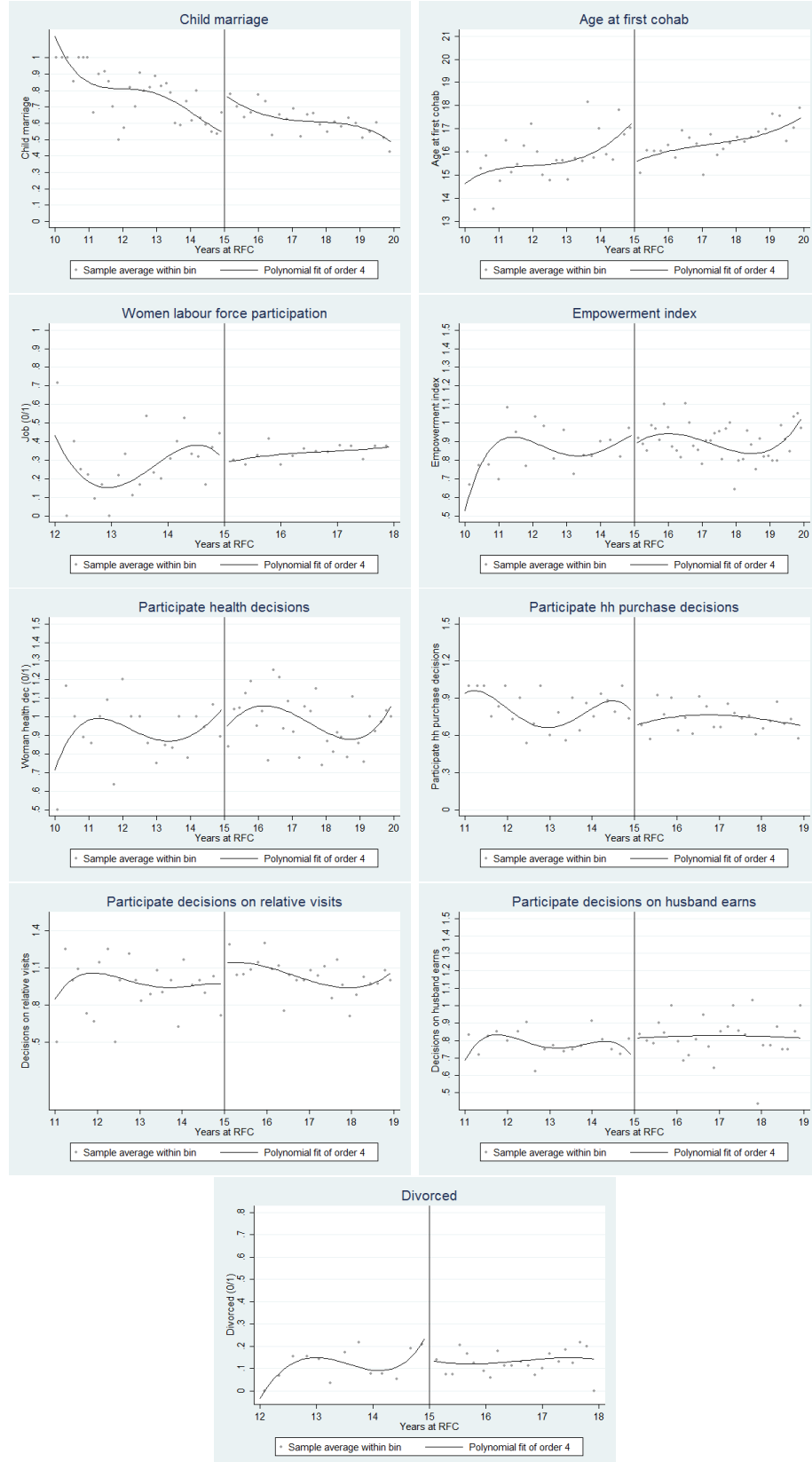
¹⁹These results are not reported in the tables provided in the study. The estimates reveal that exposure to a legal age of marriage at 18 years relative to the possibility of getting married at 15 reduces the incidence of neonatal mortality of the first born at the cut-off by 9.0 percentage points. The effect of a one-year delay in women's age at cohabitation during teenage years on the probability of neonatal mortality of the first born is estimated at 4.5 percentage points.

One important threat to the attribution of the effects identified on infant mortality to the rise in the legal age for marriage would be that the RFC not only raised the legal age for marriage but also set some additional provisions aiming to change the balance of power within the household through facilitating the procedure of divorce, abolishing the right of husbands to forbid women to work and providing women the right to administer the common marital property. These legal changes may have improved women economic status, empowerment and participation in household decisions ultimately affecting infant mortality. Although this seems a plausible possibility, the point here is that all of these norms were applied retrospectively regardless of whether women were already married or not, and therefore, they should not affect differently women aged just below and above 15 years when the RFC was approved. Nonetheless, I examine empirically whether the law affected differently labour force participation, divorce rates and participation in household decisions of women at both sides of the cut-off. To measure woman participation in household decisions, I use a set of questions in the DHS survey that provide information on who decides on relative visits, household purchases, health expenditure and the administration of the money earned by the husband. Each of these variables take the value of 0 if the woman does not participate in the decision, 1 if the woman participates in the decision and 2 if the decision is taken alone by the woman. Then, I construct a self-reported empowerment index for each woman as an average score in these questions. Using each of these self-reported empowerment measures, labour force participation and divorce rates as dependent variables, I estimate equation 1.2. The results reported in table 1.4 suggest that, overall, the new dispositions included in the RFC aiming to change the balance of power within the household did not affect differently labour force participation, divorce rates and participation in household decisions of women at both sides of the cut-off. In consequence, the evidence suggests that the effect on infant mortality identified in the study is not driven by these additional norms aiming to improve women's bargaining power within the household.

Finally, the possibility of women manipulating their reported age in the survey (the base for the construction of the forcing variable) is examined conducting a McCrary test. Figure 1.11 in appendix 1.A shows that the density of the age of women changes smoothly at the cut-off suggesting that women just below or above the cut-off age did not systematically misreport their age in the survey. Finally, the potential bias induced by measurement

error in the reported age at cohabitation in the second stage equation is addressed through the use of women's age at RFC as an instrumental variable in the regression.

Figure 1.10: RFC and women characteristics



1.7 Mechanisms

One potential mechanism driving the effect of early cohabitation on infant mortality of the first born could be the age of women at first birth. Figure 1.4 shows the strong negative association between age at first born and infant mortality among early teenage girls in the Ethiopian data. This negative association between age at first birth and infant mortality during teenage years is well documented in the medical literature and could reflect that the body of teenage women is still not optimal for the development of a successful pregnancy and/or the effect of psychological maturity on the adoption of adequate antenatal and postnatal health behaviours (Olausson et al., 2007, 1999).

To examine this mechanism, I re-estimate equations 1.1, 1.2 and 1.3 using fertility outcomes, antenatal and postnatal behaviours as dependent variables in the reduced form and second stage equations. The results on fertility outcomes are displayed in columns 1 to 4 of table 1.6 and confirm that although delaying cohabitation causally reduces the number of months between cohabitation and first birth, it also increases significantly the age of women at first birth. In line with this *age at birth* mechanism, the results reported in columns 7 and 8 suggest that the effect of age at cohabitation on infant mortality vanishes for infants given birth after the first born.

On the other hand, the effect of the age at cohabitation on the adoption of antenatal and postnatal health practices is mixed. The results reported in columns 17 to 26 indicate that delaying cohabitation increases significantly the probability of conducting a postnatal check. The coefficients measuring the effect of women's age at cohabitation on the probability of giving birth at home and on the duration of breastfeeding have the expected sign (negative the former and positive the latter) although the magnitudes are small and statistically insignificant when the optimal bandwidths are used. However, the coefficients for vaccinations and antenatal visits in the second stage equations have an unexpected negative sign, although statistically insignificant at conventional confidence levels. Nonetheless, it is important to remark that the results on the adoption of antenatal and postnatal health behaviours should only be interpreted as suggestive because DHS data only report information on these variables when the first born is alive and was born less than 5 years ago. For some of these variables, the information is only available if the child is also the last birth of the women before the survey. This could be problematic because on the one

hand the sample size used in the analysis is much smaller, reducing the statistical power of the estimations and on the other hand, the limitations in the data collection may induce a problem of sample selection bias.

Nonetheless, the age at cohabitation could also affect infant mortality of the first born through other paths. For example, a younger age of the women at cohabitation may lead to lower levels of participation in household decisions (Jensen and Thornton, 2003). Since women and men have different preferences for investment in children's health (Allendorf, 2007; Majlesi, 2014), early cohabitation may lead to higher infant mortality rates. Similarly, a younger age at cohabitation may affect investments in child's health and infant mortality through constraining women's education (Field and Ambrus, 2008) and impacting labour market outcomes (Elborgh-Woytek et al., 2013). On the other hand, given the existence of a premium for early marriage in the marriage market (Wahhaj, 2015), it is also possible that early marriage impacts infant mortality through affecting marriage market outcomes.

To investigate these paths of impact, I estimate the effect of early cohabitation on different women outcomes including participation in household decisions, marriage market outcomes, labour force participation and education attainment. The results of the estimates for equations 1.1, 1.2 and 1.3 using these outcomes as dependent variables are reported in columns 5 to 6, 9 to 16 and 27 to 28 of table 1.6. The results confirm that exposure to a legal age of marriage at 18 and delaying age at cohabitation do not seem to affect relevantly labour force participation for women, participation in household decisions, years of education and marriage market outcomes including age difference with partner, wealth index and partner's years of education.

Jointly, these results suggest that the effect on infant mortality of delaying age at cohabitation during teenage years for women seems to operate mainly through increasing the age at first birth. On the other hand and given the data limitations, we cannot rule out that the effect is also channelled through larger levels of antenatal and postnatal investments for girls cohabiting later. However, given the lack of effect of the age at cohabitation on self-reported empowerment, labour market outcomes or education, any potential effect of the age at cohabitation on infant mortality via antenatal and postnatal investments would probably be linked to an older age at first birth and a more mature behaviour leading to larger adoption of antenatal and postnatal health measures.

Table 1.6: Analysis of mechanisms

	Age at birth		Months: F. Born -F. Cohab.		Years school (compl. years)		Infant mortality Non-first born		Empowerment index		Years school partner		Wealth index	
	(1) FS	(2) RF/SS	(3) FS	(4) RF/SS	(5) FS	(6) RF/SS	(7) FS	(8) RF/SS	(9) FS	(10) RF/SS	(11) FS	(12) RF/SS	(13) FS	(14) RF/SS
Age<15 at RFC	2.070*** (0.000)	0.996*** (0.007)	2.062*** (0.000)	-10.612* (0.067)	2.064*** (0.000)	-0.706 (0.145)	2.329*** (0.000)	0.011 (0.846)	1.964*** (0.000)	-0.095 (0.192)	2.043*** (0.000)	0.783 (0.356)	2.058*** (0.000)	-0.086 (0.705)
Age at 1st cohab.		0.482** (0.013)		-5.148* (0.056)		-0.346 (0.264)		0.005 (0.860)		-0.048 (0.219)		0.384 (0.461)		-0.041 (0.754)
N		5078		5078		5078		16373		4181		4999		5078
N effect. obs.		905		935		905		737		691		864		958
Bandwidth		37.4		38.3		37.1		32.6		33.4		36.5		39.9
	Age diff. partner		Months Breastfeed.		Postnatal check		N Vaccin.		Birth at home		N antenatal visits		Work	
	(15) FS	(16) RF/SS	(17) FS	(18) RF/SS	(19) FS	(20) RF/SS	(21) FS	(22) RF/SS	(23) FS	(24) RF/SS	(25) FS	(26) RF/SS	(27) FS	(28) RF/SS
Age<15 at RFC	1.904*** (0.000)	-0.848 (0.377)	2.682*** (0.001)	-0.023 (0.997)	2.035*** (0.001)	0.268** (0.026)	0.633 (0.176)	-0.173 (0.768)	1.187*** (0.006)	-0.015 (0.880)	1.049* (0.092)	-1.016* (0.054)	2.097*** (0.000)	0.035 (0.711)
Age at 1st cohab.		-0.440 (0.444)		0.014 (0.996)		0.133* (0.080)		-0.271 (0.766)		-0.015 (0.872)		-0.981 (0.157)		0.016 (0.763)
N		4171		240		312		781		862		520		5077
N effect. obs.		853		146		204		594		584		376		828
Bandwidth		40.4		36.6		43.7		56.0		49.0		50.2		34.5

Note: Each coefficient provided in the table is estimated using a separate regression. The table reports the estimates of interest for the first stage (FS), reduced form (RF) and second stage (SS) equations using the optimal bandwidth and the robust variance estimator described in Calónico et al. (2016). The coefficients for the variable *Age<15 at RFC* measure the effect of the RFC on the age at first cohabitation (FS) and on the outcome variable analyzed (RF). The coefficients for the variable *Age at 1st cohab* measure the effect of delaying one year the age at cohabitation during teenage years on the outcome variable analyzed (SS). The sample size and the bandwidths used in the RF, FS and SS regressions are common within every outcome analyzed. The regressions conducted include as control variables a set of dummies for the regions of residence, the age of women at survey, ethnic and religion affiliation, gender of the first born, a rural/urban dummy variable and a non-parametric function for the age of the women at RFC. Columns 1-16 and 27-28 report the results for the whole sample used in the main analysis. The information on age difference with partner and on partner's years of education refer to the last partner. Column 17-26 conducts the analysis using the sample of women that had their first born within the last 5 years. For these cases, the survey provides information on maternal health outcomes. Standard errors are clustered at the forcing variable. P-values are in parentheses. ***, ***, p<0.01; **, p<0.05; *, p<0.1.

1.8 Conclusions

This study uses a regression discontinuity design to assess the impact for women of exposure to a legal age of marriage at 18 years on the incidence of child marriage, mean age at cohabitation and infant mortality of the first born. Then, the study estimates the causal effect of women's age at cohabitation during teenage years on the probability of infant mortality of the first born. The methodological design exploits age discontinuities in exposure to a law that raised the legal age of marriage for women from 15 to 18 years in some regions of Ethiopia.

The RDD estimates suggest that exposure to a legal age of marriage at 18 years relative to the possibility of marrying legally at 15 increases significantly the mean age at cohabitation by 2 years and decreases the incidence of child marriage by approximately 20 percentage points. Besides, the analysis reveals that increasing the minimum age of marriage in the law has beneficial effects on infant mortality. In the preferred estimation, exposure to a legal age of marriage at 18 relative to the possibility of getting legally married at the age of 15 decreases the probability of infant mortality of the first born child by 7.9 percentage points. The reduction in the probability of infant mortality of the first born child caused by a delay of one year in women's age at cohabitation during teenage years is estimated at 3.8 percentage points. These results are robust to the use of different estimation methods, alternative bandwidths, windows and up to 4th order spline polynomials for the forcing variable. Different placebo tests are also conducted to rule out the possibility that the effects are driven by over time reductions in infant mortality, systematic differences between women born in different months of the year, other legal dispositions included in the RFC or by national level policies affecting differently women at both sides of the cut-off. There is also no evidence of discontinuities in variables plausibly unaffected by the change in the minimum age of marriage neither of manipulation in the forcing variable at the cut-off. Finally, the results are robust to the use of neonatal mortality as dependent variable, suggesting that most of the effect of child marriage on infant mortality occurs during the first month of life.

The effects of delaying cohabitation on infant mortality are aligned with existing evidence from studies that use correlation analysis to explore the link between early marriage, teenage pregnancy and infant mortality. For example, [Raj et al. \(2010\)](#) find an odds

ratio of 1.5, statistically significant at the 95%, for the statistical association between infant mortality and women that marry before the age of 18 in India. In this line, using household data from Nepal, [Adhikari \(2003\)](#) shows that neonatal mortality rates among children of mothers 15-19 are 73% higher than among children of mothers 20-29. The effect of one-year delay in women's age at first cohabitation on the probability of infant mortality of the first born found in this study is comparable to the joint effect on child mortality at the village level of moving from 0% coverage of measles, BCG, DPT, Polio and Maternal Tetanus vaccinations to 100% ([McGovern and Canning, 2015](#))²⁰.

However, although the effects on infant mortality found in the study are large, it is important to acknowledge that the estimates yielded by the RDD employed are local and therefore, any generalization of these results to the whole population of Ethiopian women or to non-first born children should be avoided. Indeed, the analysis does not find any effect of early cohabitation on the infant mortality of children that are not the first born.

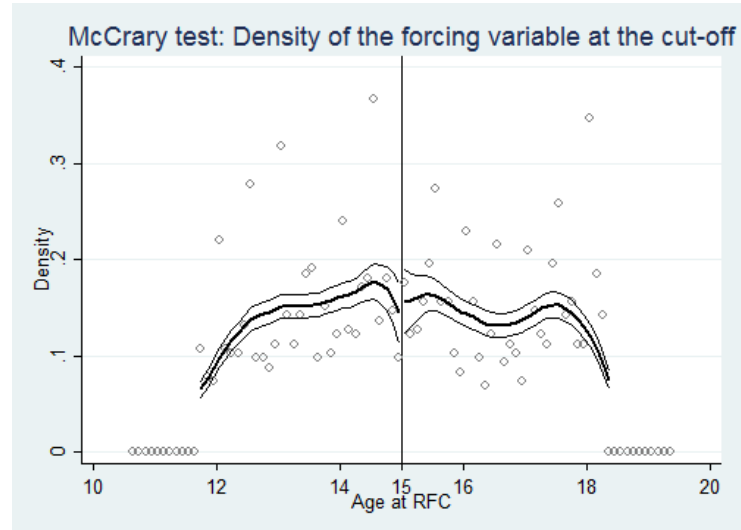
The analysis of mechanisms suggests that the strong effect of early cohabitation on infant mortality of the first born seems to be driven by the positive effect of delaying cohabitation on the age at first birth. However, due to data limitations, it is not possible to disentangle whether this *age at birth* effect is caused by purely biological reasons or by a more mature behaviour raising the adoption of antenatal and postnatal health measures. On the other hand, the analysis rules out the possibility that the effect of early cohabitation on infant mortality of the first born is driven by an effect of the former on marriage market outcomes, participation in household decisions, education or labour force participation for women.

This study contributes to the thin literature that documents the negative effects of child marriage, providing also a credible methodological alternative to previous studies relying on instrumental variable that can be used to expand the analysis to other outcomes and settings. Besides, although the effects estimated should be interpreted as local, the study shows for the first time that, even if imperfectly enforced, laws increasing the legal age of marriage can contribute to reducing infant mortality. Finally, the study provides the first causal evidence on the beneficial effect on infant mortality of delaying women's age at cohabitation.

²⁰Using DHS data from 62 countries, [McGovern and Canning \(2015\)](#) estimate at 0.73 the relative risk of child mortality at the village level associated with moving from 0% coverage of measles, BCG, DPT, Polio and Maternal Tetanus vaccinations to 100%. At the cut-off, the relative risk of the infant mortality of the first born associated with a one-year increase in the age at cohabitation is estimated at 0.75.

Appendix 1.A McCrary Test: Density of the Forcing Variable at the Cut-Off

Figure 1.11: Density of the forcing variable at the cut-off.



Appendix 1.B Additional Graphs

Figure 1.12: RFC and child marriage: Includes all women (not only those that ever cohabit and gave birth)

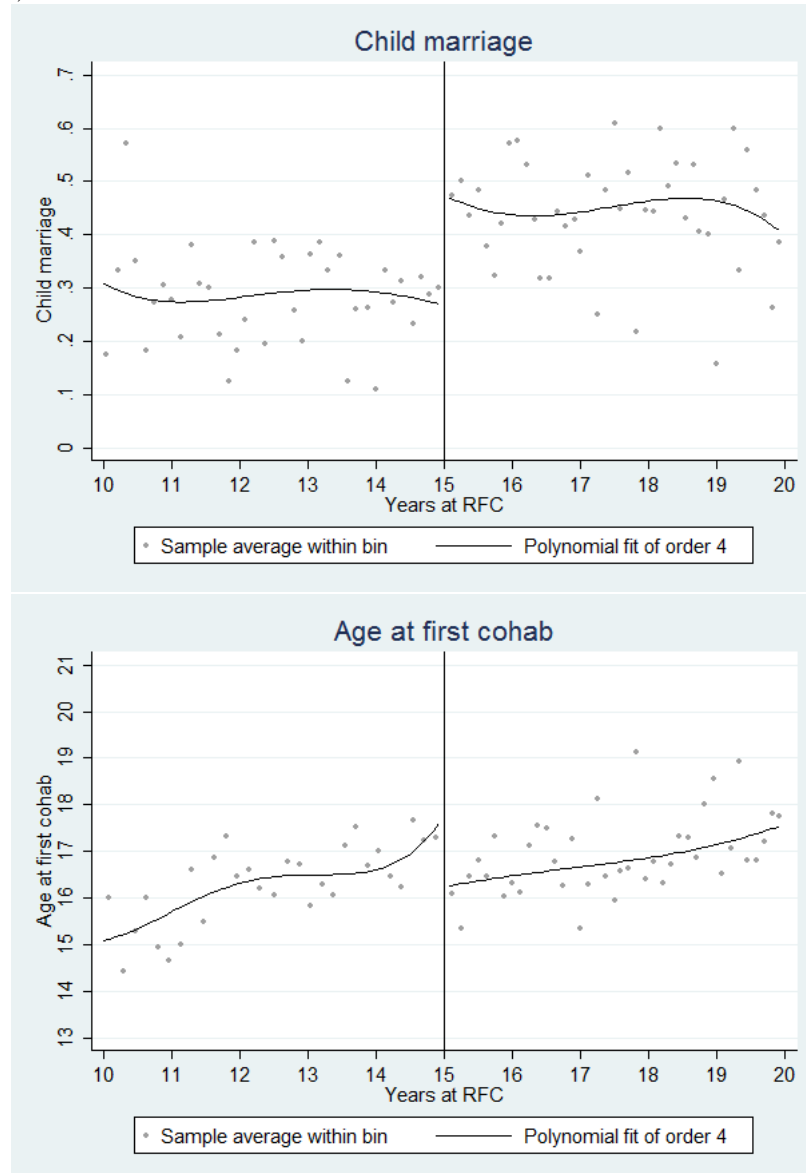


Figure 1.13: Age at first cohabitation and infant mortality of first born: Full sample of women (18-49)

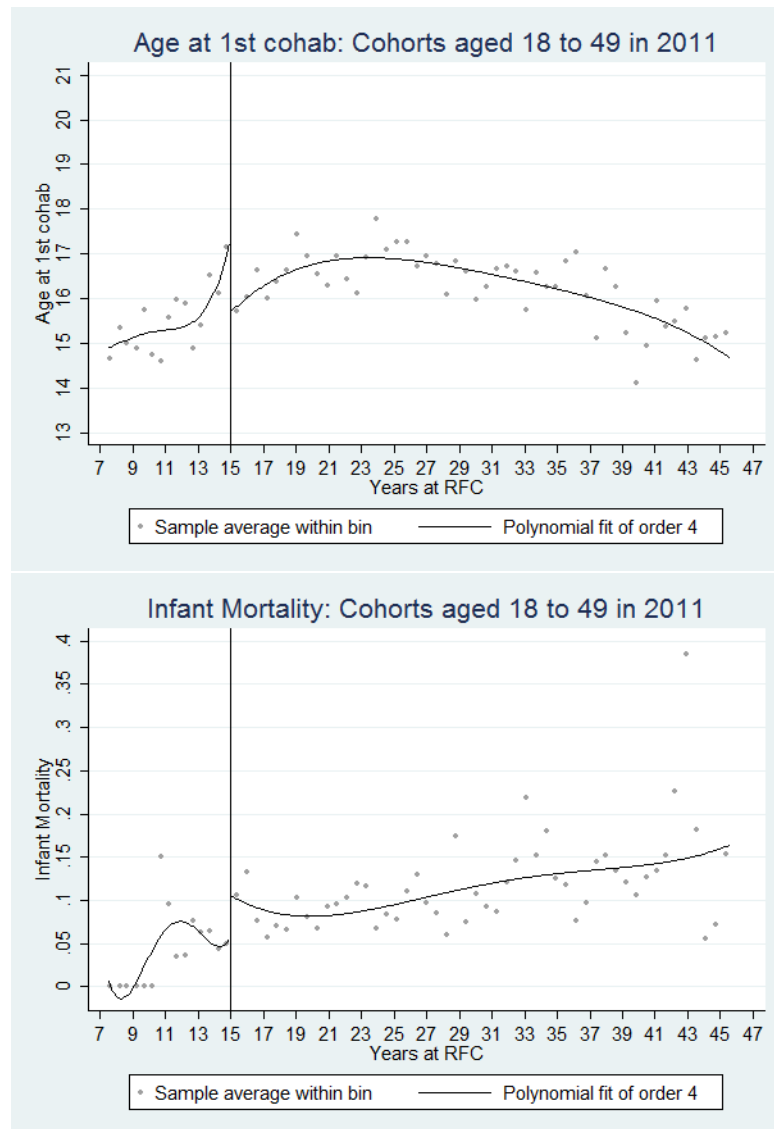


Figure 1.14: Age at 1st cohabitation by age cohort.

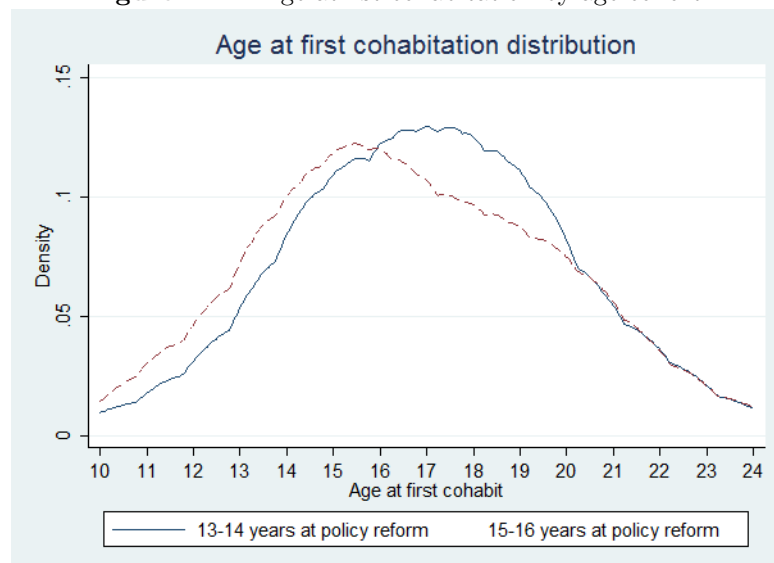


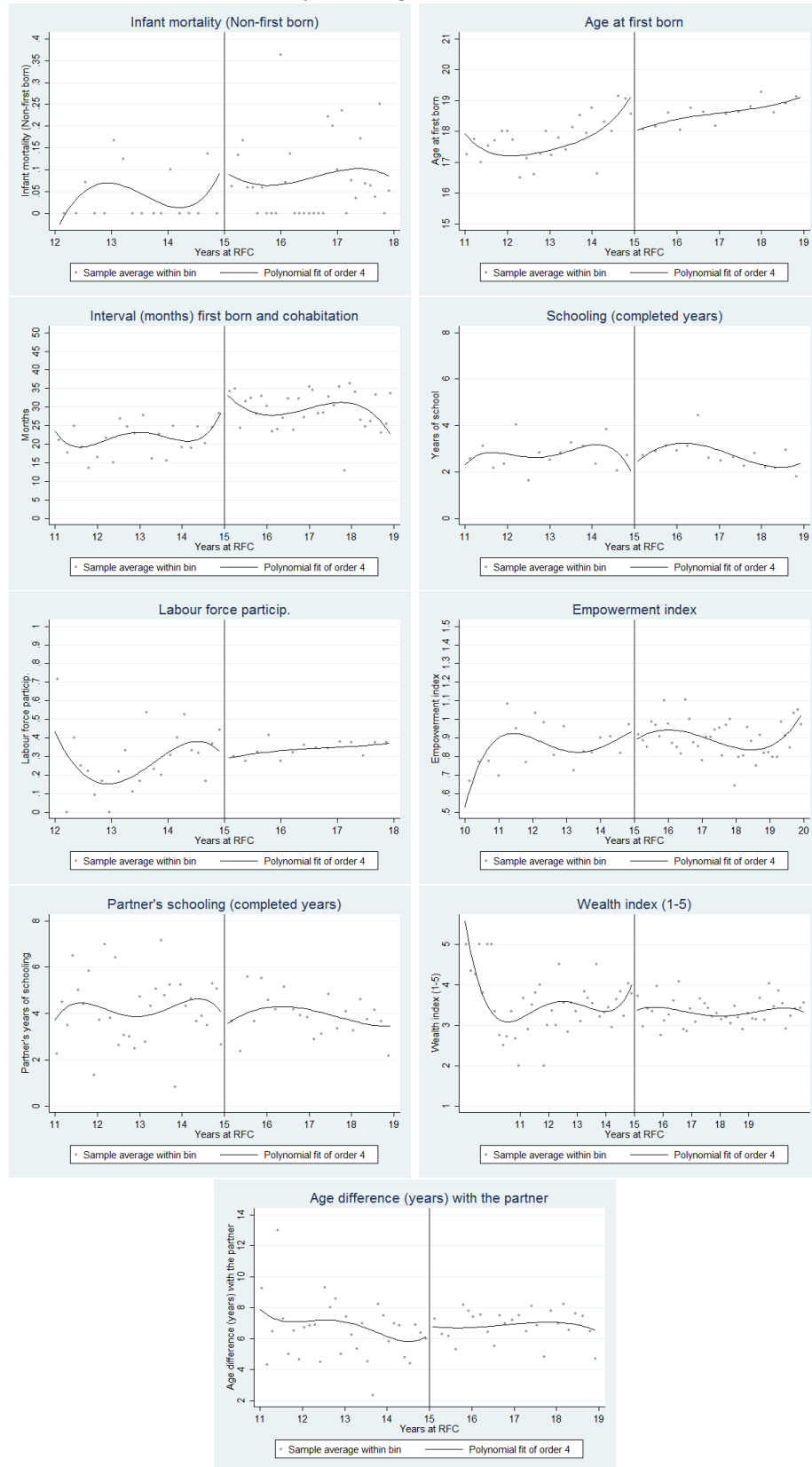
Figure 1.15: Mechanisms: Fertility, marriage market outcomes and other women outcomes

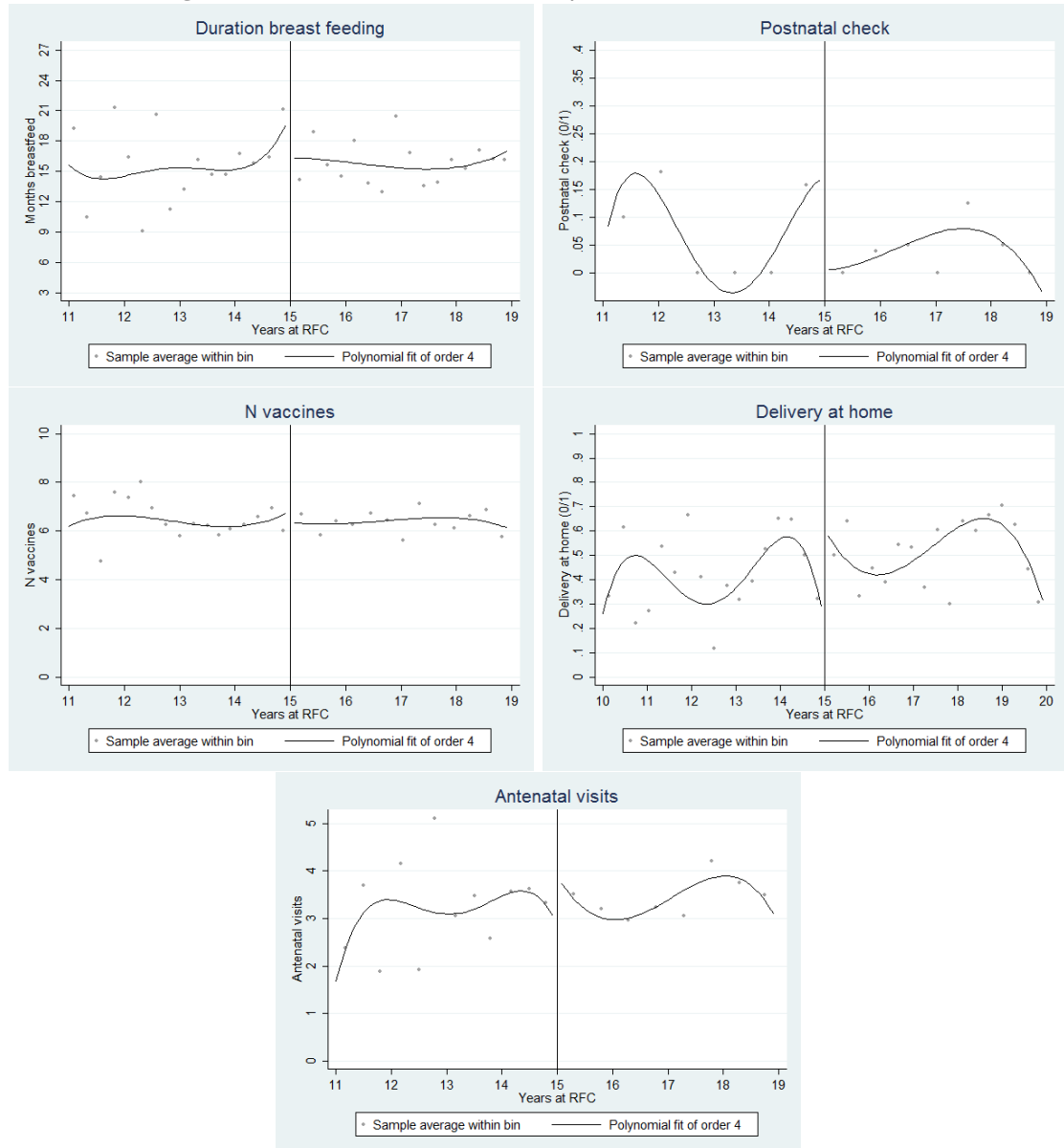
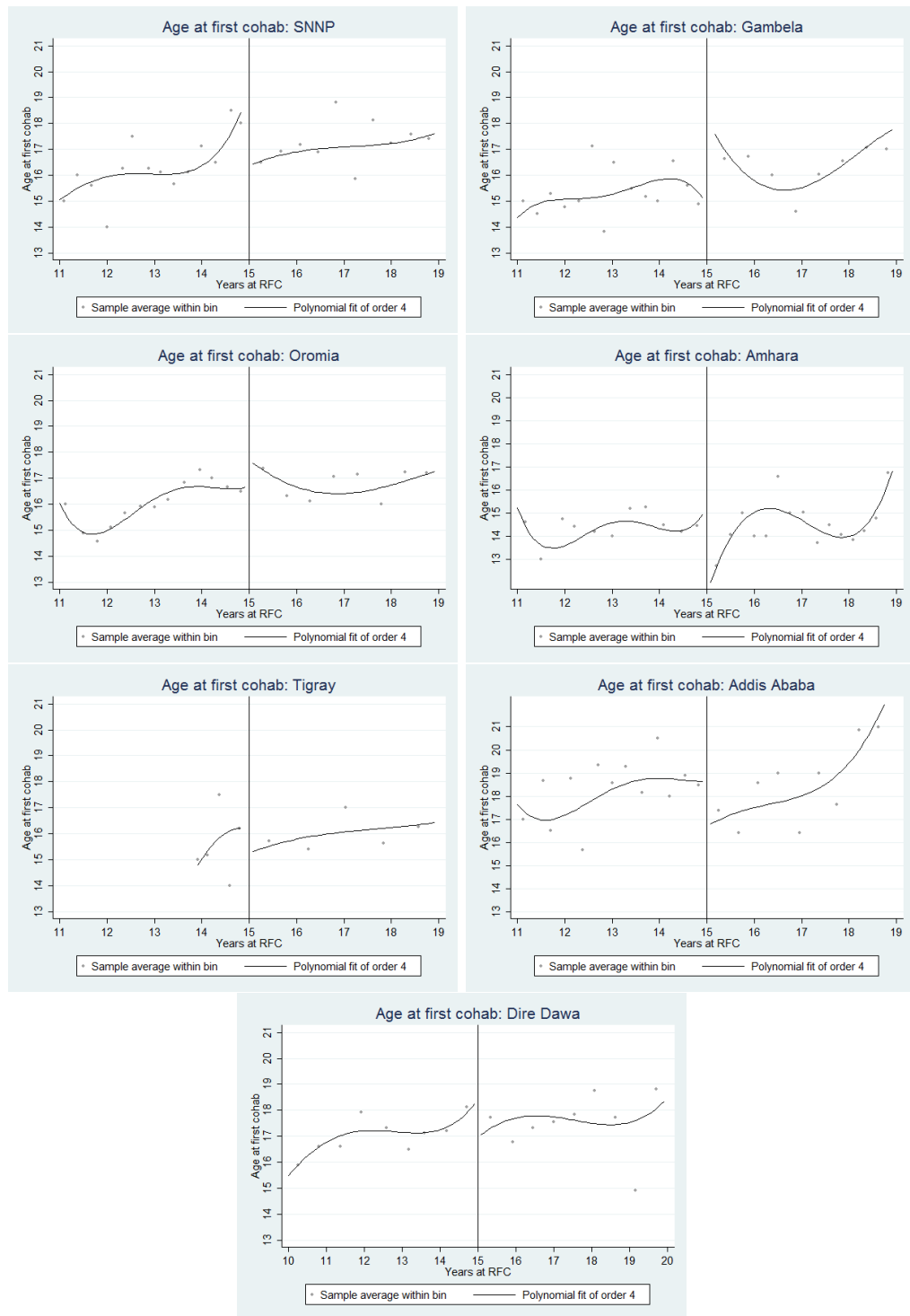
Figure 1.16: Mechanisms: Maternity and child health for first born

Figure 1.17: Age at first cohabitation at the cut-off for the Ethiopian regions that approved RFC between 2000 and 2007



Appendix 1.C Sample of Women Still Cohabiting with the First Partner

Table 1.7: Only women still cohabiting with first cohabit.

	Age at birth		Infant Mortality		Years school		Anaemia	
	(1) FS	(2) RF/SS	(3) FS	(4) RF/SS	(5) FS	(6) RF/SS	(7) FS	(8) RF/SS
Age<15 at RFC	1.233*** (0.005)	1.204*** (0.001)	1.291*** (0.003)	-0.097* (0.085)	1.324*** (0.002)	-1.020 (0.260)	1.268*** (0.003)	-0.094 (0.228)
Age at 1st cohab.		0.977*** (0.003)		-0.076* (0.082)		-0.766 (0.420)		-0.075 (0.351)
N		3126		3126		3126		2985
N effect. obs.		595		651		651		739
Bandwidth		34.8		37.9		37.1		43.0

	Empowerment index		Years school partner		Age Difference		Work	
	(9) FS	(10) RF/SS	(11) FS	(12) RF/SS	(13) FS	(14) RF/SS	(15) FS	(16) RF/SS
Age<15 at RFC	1.270*** (0.004)	-0.116* (0.098)	1.286*** (0.003)	-0.028 (0.981)	1.336*** (0.002)	-0.964 (0.293)	1.258*** (0.004)	-0.014 (0.861)
Age at first cohabit.		-0.091 (0.156)		-0.022 (0.984)		-0.692 (0.349)		-0.011 (0.874)
N		3118		3102		3112		3125
N effect. obs.		593		647		691		626
Bandwidth		34.4		37.9		39.2		36.5

Note: The analysis reported in the table is conducted using the sample of women 18-49 that ever bore a child and still cohabit with first partner. Each coefficient provided in the table is estimated using a separate regression. The table reports the estimates of interest for the first stage (FS), reduced form (RF) and second stage (SS) equations using the optimal bandwidth and the robust variance estimator described in [Calonico et al. \(2016\)](#). The coefficients for the variable *Age<15 at RFC* measure the effect of the RFC on the age at first cohabitation (FS) and on the outcome variable analyzed (RF). The coefficients for the variable *Age at 1st cohab* measure the effect of delaying one year the age at cohabitation during teenage years on the outcome variable analyzed (SS). The sample size and the bandwidths used in the RF, FS and SS regressions are common within every outcome analyzed. The regressions conducted include as control variables a set of dummies for the regions of residence, the age of women at survey, ethnic and religion affiliation, gender of the first born, a rural/urban dummy variable and a non-parametric function for the age of the women at RFC. Standard errors are clustered at the forcing variable. P-values are in parentheses. ***p<0.01; **p<0.05; *p<0.1.

Chapter 2

Cognitive Skills and Intra-Household Allocation of Schooling

2.1 Introduction

Understanding how households allocate resources among their members has been an inquiry of primary interest for economists for decades ([Chiappori and Meghir, 2014](#)), with important implications for the intergenerational transmission of human capital ([Datar et al., 2010](#)). One key aspect of the intra-household allocation of resources is the role of cognitive skills in the distribution of school investments across the children of a household. Do parents allocate more schooling to more able children or do they try to compensate less skilled children with more schooling? Answering this question is of crucial importance for the design of effective policies that pursue improvements in education for all the population. For example, if parents focus their school investments in their most endowed children, supply-side schooling interventions such as reductions in class size or schools construction may benefit mostly the most able children while demand-side schooling interventions that target the less able children such as conditional cash transfers could be more effective in promoting universal schooling ([Akresh et al., 2012](#)).

The winner of the 1992 Nobel Prize in Economics Gary Becker proposes in his seminal book *A Treatise on the Family* ([Becker, 1981](#)) a theoretical model that conceptualizes the intra-household allocation of human and nonhuman capital investments across siblings. One of the main predictions of the model is that siblings with higher returns to human capital receive larger human capital investments. Assuming that returns to human capital

investments are larger the higher the cognitive ability (Appleton, 2000; Becker, 1981), the model predicts that parents reinforce genetic differences in cognitive skills through allocating more human capital investments to more able siblings and compensate less endowed siblings with more nonhuman capital investments.

Using longitudinal data from rural Ghana, this study builds in the existing literature and tests one of the predictions of Becker’s model: Do parents allocate investments in education reinforcing cognitive differences between siblings? More specifically, the study assesses whether in a context in which the intra-household variation in school attendance across siblings in compulsory and post-compulsory school age is large, better cognitive skills relative to the rest of the siblings in the household affect the probability of attending school. This chapter makes three contributions. First, the study adds evidence to the thin literature that examines whether cognitive skills affect the intra-household allocation of school investments. Second, to achieve this objective, the study uses 4 rounds of a panel dataset from rural Ghana. The empirical strategy relies on the use of cognitive tests applied in the first round of the survey as *treatment* variables to proxy for cognitive skills, and school attendance the subsequent years as dependent variable in the specification. Variables aiming to capture human capital investments before the implementation of the cognitive tests are included as control variables in the specification to address endogeneity concerns. Although this empirical strategy is not without its limitations, I show that it represents an advantage relative to previous papers investigating the same question. Finally, this is the first study testing whether the effect of cognitive skills on the allocation of schooling across siblings depends on the gender of the child, household wealth, household size or on whether the household is polygynous.

I find suggestive evidence that parents reinforce cognitive skills differences through allocating more schooling to those children that score better in cognitive skills tests. These results are aligned with previous papers that assess empirically the role played by cognitive skills in the intra-household allocation of investments in education (Kim, 2005; Akresh et al., 2012; Ayalew, 2005) and provide support for the main prediction of Becker’s model. Furthermore, the effect of cognitive skills on the allocation of schooling across siblings seems larger for boys than for girls although the evidence is not conclusive. On the other hand, the results also suggest that the magnitude of this effect does not depend on the household characteristics analysed.

The paper is structured as follows. Section 2.2 presents the conceptual framework. Then, section 2.3 summarizes the literature that investigates the role played by child endowments in the allocation of resources across siblings. Section 2.4 introduces some relevant aspects of the educational system in Ghana. Section 2.5 describes the data. Section 2.6 presents the empirical strategy and section 2.7 discusses the main results of the study. Then, section 2.8 examines whether the effect of cognitive skills on the allocation of schooling across siblings varies across different types of households and depends on the gender of the child. Section 2.9 concludes.

2.2 Conceptual framework: Becker (1981)

This section summarizes the theoretical model developed in Becker (1981)¹. This model aims to formalize the intra-household allocation of human and nonhuman capital across the siblings of a household. In it, parents maximize a utility function that depends on their own current consumption and on the future wealth of their children:

$$U = U(c, I_1, \dots, I_n) \quad (2.1)$$

where c is the consumption of the parents in the present and I_i indicates the future wealth of child i . The future wealth of a child is described as a function of the human capital and nonhuman capital investments received by this child:

$$I_i = R_i^h(h_i, a_i)h_i + R^m m_i \quad (2.2)$$

where h_i and m_i indicate the level of human and nonhuman capital resources that parents allocate to child i . R_i^h is the rate of return on human capital for child i , a_i indicates the cognitive skills of the same child, and R^m is the market rate of return on nonhuman capital. The model assumes that while the rate of return on human capital function is concave in the investment level and higher the larger the innate cognitive skills, the rate of return on nonhuman capital is constant in the investment level and independent of innate cognitive skills implying that it is the same across siblings. Formally, if $R_i^h(h_i, a_i)$ and R_i^m are the functions of marginal returns on human and nonhuman capital, Becker's model assumes that $\delta R_i^h / \delta h_i < 0$, $\delta R_i^h / \delta a_i > 0$, $\delta R_i^m / \delta m_i = 0$ and $\delta R_i^m / \delta a_i = 0$. The

¹The notation used is based on Kim (2005).

assumption made in Becker's model of a larger marginal rate of return on human capital for better endowed siblings has some empirical support in the literature ([Appleton, 2000](#)).

Parents maximize their utility function through allocating human and nonhuman capital resources until the marginal rate of return on human capital is equal across siblings and also equal to the rate of return on nonhuman capital. Formally, the latter condition can be expressed as follows: $R_1^h = R_2^h = \dots = R_n^h = R^m$.

The model yields two main predictions. First, because they have higher returns on human capital, more able siblings receive larger levels of human capital investments than less cognitively endowed siblings. In consequence, with their human capital investments, parents reinforce cognitive differences between siblings. This prediction is straightforward in the model: if under optimal levels of investment across siblings $R_1^h(h_1, a_1) = \dots = R_n^h(h_n, a_n) = R^m$, $\delta R_i^h / \delta h_i < 0$, $\delta R_i^h / \delta a_i > 0$ and for the following two siblings we assume $a_1 > a_2$, then $h_1 > h_2$. More generally, the model implies that $\delta h / \delta a > 0$.

The second prediction of the model is that parents compensate siblings with lower returns on human capital through allocating them larger bequests or other nonhuman capital resources (e.g. inter-vivos transfers unrelated with education or health). The implication of this prediction is that less able siblings receive larger nonhuman capital investments. To see how the model leads to this conclusion, I use as example a household with two siblings where $a_1 > a_2$. In this household, parents maximize their utility when $\frac{\delta U}{\delta I_1} \bigg/ \frac{\delta U}{\delta I_2} = \frac{R_2}{R_1}$, where R_i is the rate of return from additional investments in child i . Because in the optimal level of investments $R_1 = R_2 = R^m$, the former condition can only be satisfied when $I_1 = I_2$. Thus, if $a_1 > a_2$ and $h_1 > h_2$, the condition $I_1 = I_2$ can only be achieved if $m_1 < m_2$.

Using a sample of Ghanaian children, the core of this study focuses on testing the first prediction of the model using years of schooling as a measure of human capital investments and the scores in different cognitive tests as proxies for cognitive skills. Furthermore, using the conceptual framework described above as a departing point, section 2.8 examines whether the magnitude of the effect of cognitive skills on the allocation of schooling across siblings depends on the gender of the child and on household level characteristics such as household size, polygyny status or household wealth. Unfortunately, the lack of information on the intra-household allocation of nonhuman capital resources precludes the empirical assessment of the second prediction of the model.

2.3 Related Literature

Several studies assess empirically the effects of child endowments on the allocation of human and nonhuman capital resources across siblings, testing the predictions of the Becker’s model presented in section 2.2. Most of these studies focus on examining the effect of health endowments such as birthweight or body mass index (BMI) on the intra-household allocation of health, schooling and other inter-vivos transfers. The evidence is mixed and while some of these studies find that children with better health seem to receive larger levels of health (Jere R. Behrman, 1994; Pitt et al., 1990) and educational investments (Jere R. Behrman, 1994; Paul Miller, 1995), other studies suggest that parents favour less endowed children with more education (Frijters et al., 2010; Griliches, 1979; Jere R. Behrman, 1982; Ermisch and Francesconi, 2000; Leight, 2014) and health investments (Datar et al., 2010; Ayalew, 2005). The evidence is also inconclusive on whether children with worse health receive more nonhuman capital transfers, including inter-vivos transfers unrelated with health or education (McGarry and Schoeni, 1995; Dunn and Phillips, 1997; Hochguertel and Ohlsson, 2009; Wolff, 2006) .

Although plenty of studies assess whether child endowments affect the intra-household allocation of human capital investments, only a few of them investigate the specific role of cognitive endowments. Indeed, there are only three studies that use data on cognitive tests to explore empirically the role of cognitive endowments in the allocation of human capital investments across the children of the household. The main methodological challenge in these studies is the potential endogeneity in the link between the results in cognitive tests such as IQ, Raven or Digit Span tests obtained by the child and the human capital investments that this child has received.

The first study that addressed empirically this research question was Kim (2005). Using data from high schools in the state of Wisconsin, in the US, the author shows that conditional on a wide set of household and individual level controls, higher scores in IQ tests are associated not only with receiving larger human capital investments but also with receiving larger inter-vivos transfers unrelated with education. This finding provides evidence in favour of the first prediction of Becker’s model but against the second. Nonetheless, we cannot rule out the possibility that the statistical association between IQ and investments received is partially driven by reverse causality.

Using cross-sectional data from Ethiopia, [Ayalew \(2005\)](#) finds that while parents allocate more health investments to children with worse cognitive abilities, they also tend to reinforce cognitive differences through allocating more educational investments to those children that have better cognitive skills. To measure these skills, the author constructs a proxy variable for innate cognitive ability as follows. First, the score obtained in the Raven cognitive test is regressed against individual characteristics of the child including age, gender and years of schooling. Then, the residuals in the latter regression are used as a proxy for innate cognitive ability in a main specification that includes school attendance as dependent variable and the residuals of the Raven score regression as the explanatory variable. The main problem of this empirical strategy is that, in order to overcome the endogeneity in the link between years of schooling and the Raven score in the first regression, years of schooling are instrumented with land and livestock owned. This solution would violate the instrumental variables exclusion restriction if land or livestock owned affect the performance in the Raven test through any mechanism other than through years of schooling. If so, the estimated effect of innate ability on school attendance in the main regression would also be biased.

The most recent paper that investigates the effect of cognitive skills on the allocation of schooling investments across siblings is [Akresh et al. \(2012\)](#). Relying on panel data from Burkina Faso, the authors show that higher scores in cognitive tests and better parents' expectations regarding future earnings relative to the rest of the siblings in the household, increase significantly school attendance. The study uses the results of Raven, Forward Digit Span and Backward Digit Span tests as direct measures of cognitive skills. Then, to test whether the results are driven by reverse causality, the authors exploit the panel data dimension of the data restricting the analysis to (a) children aged 5 to 7 years old and therefore, not yet enrolled in school at the time of the collection of the first wave of surveys, and to (b) children in grades 1 and 2, for whom the authors argue that the score in the cognitive tests is not significantly affected by school attendance.

Remarkably, the results of these three studies in terms of the effect of cognitive skills on the allocation of investments in education across siblings are in line with the first prediction of [Becker \(1981\)](#). On the other hand, both [Kim \(2005\)](#) and [Ayalew \(2005\)](#) fail to find any evidence of reinforcing mechanisms in terms of other human capital investments (e.g. health investments) or compensatory behaviour using nonhuman capital transfers.

In the light of the existing evidence, some studies suggest that differences in parental investments across their children might not be primarily caused by larger returns on investments in some of the children in the household (Kim, 2005; Mechoulam and Wolff, 2015). For example, Kim (2005) suggests that the allocation of human and nonhuman capital resources across siblings could be mainly driven by parental differences in affection for their children.

2.4 Education in Ghana

When Ghana achieved its independence in 1957, the vast majority of its citizens lacked access to education (MacBeath, 2010). The post-colonial government implemented different programmes with the objective of increasing access to primary and secondary education across the country and although the process faced important challenges such as the lack of qualified teachers, the enrolment rates raised dramatically in all the country (MacBeath, 2010; Addy, 2013). Over the following decades, a free and compulsory educational system was consolidated in Ghana, with average levels of enrolment and gender parity above the average for Sub-Saharan countries (UNESCO, 2014; USAID, 2009). World Bank data from 2013 show that while the average net enrolment rates in primary and secondary education in sub-Saharan Africa were 77% and 33%, the net enrolment rates in primary and secondary education in Ghana were 87% and 52%. Furthermore, and unlike other Sub-Saharan countries, the net enrolment rates in primary and secondary education in Ghana are not significantly different for boys and girls (GSS, 2015).

However, despite achieving relatively large average enrolment rates and gender parity, the education system in Ghana is facing important challenges. An analysis of some of them is provided in UNESCO (2014). Together with the lack of infrastructure and delayed attendance, two threats gather the attention from policy makers and international organizations working in West Africa. First, different reports highlight that the quality of primary and secondary education is deficient in many schools in both urban and rural areas². Indeed, Ghana took last position in the 2015 OECD Global Education ranking in the category *Math and Science*. This report ranked internationally the quality of education for different subjects in 76 countries³. Second, Addy (2013) shows that

²See for example the UNICEF country report available at: http://www.unicef.org/about/annualreport/files/Ghana_COAR_2010.pdf

³see <http://www.bbc.co.uk/news/business-32608772>

school attendance remains low in the less-developed areas of the country, particularly in the north of Ghana, where average enrolment rates in primary and secondary education are significantly lower than national averages. For example, according to the 2014 Ghana Demographic and Health Survey, the net enrolment rates in primary and secondary school in the regions where the villages sampled are located were 68.1 and 30.0 in the Northern Region and 73.1 and 30.4 in the Upper East Region (GSS, 2015). Interestingly, girls in these regions are slightly more likely to be attending primary school than boys.

After the 2009 reform, formal education in Ghana is structured in three different parts. The basic education starts at age 4 and finishes at 15, and it is free and compulsory. During the basic education cycle, the students follow 2 years of pre-school education, 6 years of primary education and 3 years of junior secondary school. In this education cycle, grade promotion is automatic if the child attends school regularly during the year. At the end of the third year of the junior secondary school, the students are eligible to take the Basic Education Certificate Examination (BECE). Passing the latter exam gives students access to senior secondary education or vocational education in Ghana. Both the senior secondary education and vocational education last for three years and the state provides them for free in public schools. However they are not compulsory and promotion at the end of the year is based on school performance. At the end of secondary school, the students are eligible to take the West African Senior School Certificate Examination, which is required to access tertiary education. Although the duration of tertiary education depends on the academic degree undertaken, bachelor degrees in public universities typically last for four years and most students face small tuition fees.

2.5 Data

The analysis conducted in this study relies on four rounds of a household panel survey conducted annually between 2012 and 2015 for the impact evaluation of the Millennium Villages Project in the districts of Builsa and West Mamprusi in northern Ghana. The survey targets every year the same 2250 households in 103 villages in the districts of Builsa and West Mamprusi. During this period, a total of 2,080 households were successfully interviewed every year.

With the first round of the survey, implemented in 2012, the enumerators applied

three widely used cognitive tests aiming to measure cognitive skills of every person aged 5-19 in the sample regardless of whether this person ever attended school: the Raven's Progressive Matrices, the Digit Span Forward (DSF) and the Digit Span Backward (DSB) tests. The Raven's Progressive Matrices is a nonverbal test developed by John C. Raven in 1936 that measures abstract reasoning and has been used to measure fluid intelligence. The test administered includes 12 questions and the score ranges between 0 and 12. In these questions, the children have to identify the missing element that completes patterns of increasing difficulty until they get an answer wrong⁴. The test was designed to be unambiguously interpretable and easy to administer (Raven et al., 1994). Although the test was constructed to be affected as little as possible by education and experience, some studies suggest that these two features may influence test performance. For example, Ayalew (2005) and Alderman (1995) show that boys in rural areas in Pakistan and Ethiopia obtain significantly higher scores in the Raven test than girls, casting doubts on whether the score is influenced by human capital investments.

The DSF and DSB tests are components of the Wechsler memory scales (WMS) and the Wechsler intelligence scales for adults and children (Woods et al., 2011). In these tests, the children have to repeat different strings of numbers both in the order stated by the enumerator (forward test) or in reverse order (backward test). The strings of numbers increase in length as the child answers correctly up to a total of 8 digits and the score ranges between 0 and 16. These tests measure working memory and ability to concentrate⁵. Following Akresh et al. (2012), I compute age-adjusted z-scores⁶ for each of these cognitive tests so that the mean and the standard deviation of the score in each test for children of the same age in completed years are 0 and 1.

A crucial stage of the study is the construction of variables that measure investments in education. Given that the study is set in deprived areas of a low-income country with large intra-household variation in terms of school attendance, I follow Akresh et al. (2012) and use school attendance as a measure of school investments. Information on school attendance is collected yearly at the individual level in the survey during the period 2012-2015. The survey includes two different questions aiming to record school attendance.

⁴ An example of a question in the Raven Matrices test is provided in appendix 2.A.

⁵ Examples of questions in the Digit Span Forward and Backward tests are provided in appendix 2.A.

⁶ Age-adjusted z-score are computed through subtracting to every cognitive test score the mean score of the same test for the children with the same age in completed years and then dividing by the standard deviation of the score for these children.

First, enumerators collect information on years of schooling for every child equal or older than 3 years old. Using this information, I construct a measure of school attendance between 2013 and 2015 as the difference in years of schooling reported in 2015 and in 2012. Second, the survey also asks whether children equal or older than 3 years old attended school at any point during the previous 12 months. I use this information to construct a variable that ranges between 0 and 3 and measures how many years during the period 2013-2015 the individual has attended school at least once during the last 12 months.

The main concern with the second attendance measure is that it might not truly capture *relevant* attendance because children that for example attend school for just one week during the year might be categorized as attending school using this criteria. Indeed, the percentage of children aged 5-18 that attended school at least once during the last twelve months remains above 70% for children aged 5-18 and over 60% for individuals aged between 19-21. Furthermore, since only 27% of the households with more than 1 child in school age present between-siblings variation in terms of school attendance at some point in the last 12 months, this measure of school attendance is not ideal to conduct the analysis proposed. On the other hand, the main concern with the first attendance measure is that it is constructed using years of schooling rather than attendance, which might not be driven only by school attendance but also by the ability of the individual to promote to the next school grade. Although this concern could be relevant for children undertaking senior secondary school, tertiary education or vocational education where access and grade promotion at the end of the school year is not automatic, the share of children in the sample attending any of these educational levels at baseline is below 2%.

In the light of these facts and given the intermittent school attendance of most children in the sample⁷, the school attendance variable based on the difference in years of schooling reported in 2012 and in 2015 seems to be more adequate to measure investments in education than the school attendance variable based on whether a child attended school at least once in the last 12 months. Nonetheless, and although the main analysis uses the measure of school attendance based on the difference between years of schooling reported in 2015 and in 2012 as the main dependent variable, I also test the robustness of the results to the use of the alternative school attendance measure mentioned above.

⁷The prevalence of intermittent school attendance is discussed at the end of paragraph 10 in section 2.5.

The sample of primary interest for the analysis are those children aged 5-18 that are sons or daughters of the household head and live in a household interviewed every year over the period 2012-2015 with at least two siblings aged 5-18 when the first round of the survey was implemented in 2012. In total, 4,003 children from 1,468 households fulfil these conditions. However, and despite all these children were eligible for the cognitive tests, only 2,489 of these children from 1,010 households took at least one of the three cognitive tests. The majority of children that did not take any cognitive test were not at home at the time of the interview, accounting for 62-71% of the eligible children that did not take the test, depending on the cognitive test examined. The rest of these children were at home but decided not to take any cognitive test.

Table 2.1: Summary statistics: Individual characteristics of children aged 5-18 in 2012.

	Took cognitive test		Did not take cognitive test		
	Mean	Standard deviation	Mean	Standard deviation	Difference
<i>Individual level characteristics</i>					
Age	10.29	3.74	11.53	4.30	-1.23***
Female	0.46	0.50	0.47	0.50	-0.02
Sibling rank	3.41	2.20	3.67	2.65	-0.26**
Health insurance scheme	0.72	0.45	0.65	0.48	0.07***
Ever attend. school	0.85	0.36	0.74	0.44	0.11***
Attendance 2012	0.81	0.39	0.67	0.47	0.14***
N years attended 2013-2015	2.47	1.00	2.01	1.26	0.47***
Years of school 2012	3.59	2.89	3.77	3.63	-0.18
Years of school 2015-2012	1.51	1.20	1.36	1.26	0.16***
Paid work	0.20	0.40	0.30	0.46	-0.10***
Raven test results (1-12)	4.09	2.14			
Raven age-adjusted z-score	-0.01	1.00			
Digit span forward results (1-16)	4.42	3.00			
DSF age-adjusted z-score	0.00	1.00			
Digit span backward results (1-16)	1.75	1.91			
DSB age-adjusted z-score	0.01	1.01			
N of households	1,010		765		
N of children	2,489		1,514		

Note: Values of individual level characteristics based on the first round of the survey (2012). The number of households that corresponds to the 1,514 children eligible to take the test that did not take it is 765 rather than 458 because some of these 1,514 children are from some of the 1,010 households where at least one child took a cognitive test. ***p<0.01; **p<0.05; *p<0.1.

Table 2.1 provides descriptive statistics on child level characteristics for the sample of interest. The statistics are presented separately for children that took the cognitive tests and for children that did not. The average number of correct answers in the sample of

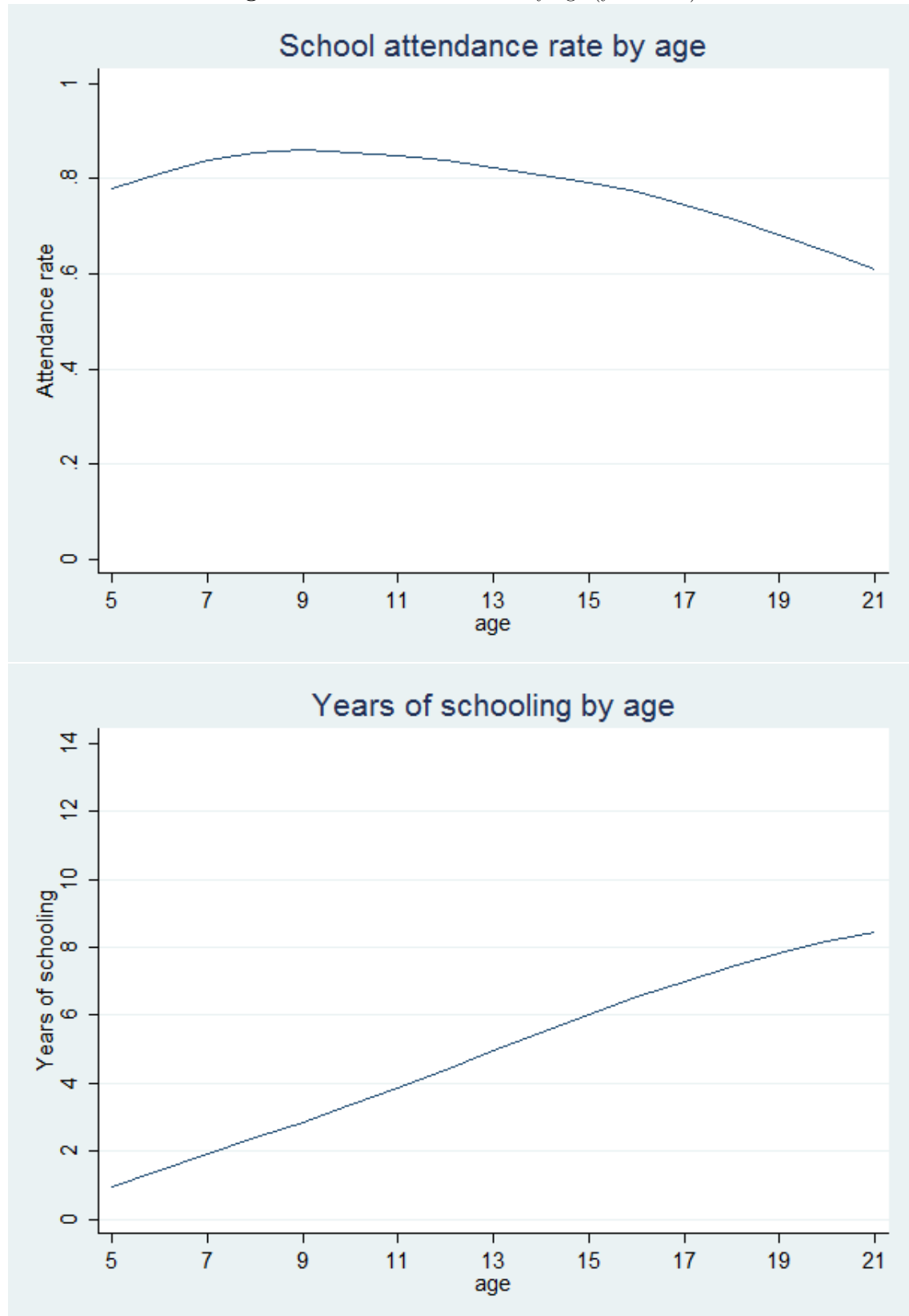
children that took the cognitive tests was 4.09 for the Raven Matrices test, 4.42 for the DSF test and 1.75 for the DSB test. The share of girls in this sample is 0.46 and the mean age of these children is 10.29 years.

The data reveal that 85% of the children aged 5-18 in 2012 that took the cognitive tests ever attended school and that 81% of these children attended school at least once in the last 12 months. Consistently, the table shows that the average number of years during the period 2013-2015 that these children attended school at least once is 2.47. The corresponding net attendance rate to primary and secondary school among this sub-sample of children in the data is 53% and 23.4%⁸. The net enrolment rates found in the sample are smaller than the net enrolment rates reported in [GSS \(2015\)](#) for the Northern and the Upper East regions for primary (68.1% and 73.1%) and secondary education (30.0% and 30.4%). When attendance during the same period is measured using the difference between years of schooling in 2015 and 2013, the average number of years of schooling completed during the relevant period for the sample of children that took at least one of the cognitive tests is 1.51.

Figure 2.1 provides further insights into school attendance and years of schooling by age for the sample of children that took at least one of the cognitive tests. The figure reveals that the share of individuals aged 5-21 that attended school at some point in the last twelve months remains over 70% until the age of 18 and over 60% at the age of 21. Interestingly, the share of children attending school decreases slightly but constantly from the age of 10. Besides, the figure shows that the slope of the LOWESS regression that displays the relation between the age and the years of schooling is well below 1 during all the period of interest. Combined with the fact that the percentage of children that reports having attended school at least once in the last 12 months remains above the 80% for children younger than 16 years old, the small slope in every point of the function suggests that despite attending school at some point during the year, many of these children might not be attending school regularly during the year.

Worth to mention, table 2.1 also shows that children that did not take the cognitive tests are on average 14 percentage points less likely to have attended school at least once during the last 12 months, 1.2 years older, 10 percentage points more likely to have a job,

⁸Primary education includes the 6 years of primary education, excluding the 2 years of pre-primary education. The group of reference for the calculation of the net enrolment rate in primary education are children aged 6 to 12 years old. The secondary education includes the 3 years of junior secondary school and the 3 years of senior secondary school. The group of reference are the children aged 13 to 18 years old.

Figure 2.1: School attendance by age (year 2012)

7 percentage points less likely to have a health insurance and have a higher sibling rank. On the other hand, there is not any significant difference between children that did and did not take cognitive tests in terms of gender, years of schooling at baseline and years of schooling completed between 2012 and 2015.

Table 2.2 presents descriptive statistics on household characteristics collected in the first round of the survey (2012) for the sample of households surveyed every year that have at least 2 children that were eligible to undertake the cognitive tests. The information is provided separately for those households in which at least one child took one or more cognitive tests and for those in which despite having eligible children, none of them took any of the cognitive tests. The descriptive statistics reveal that 88% of the households in the sample with at least one child that took the cognitive tests live on a level of per capita consumption below the national poverty line, 31% of them are polygynous and only 9% of them are female headed. The average household size in the sample of interest is 8.18 and the average number of household members aged below 18 is 4.64.

Table 2.2: Summary statistics: Household characteristics of children aged 5-18 years in 2012.

	HH took cognitive tests		HH did not take cognitive tests		
	Mean	Standard deviation	Mean	Standard deviation	Difference
<i>Household level characteristics</i>					
P/c expenditure (cedis)	742.37	654.74	723.25	618.73	19.12
Poor	0.88	0.33	0.90	0.30	-0.02
Economic shock index (0-5)	3.46	1.31	3.47	1.37	-0.01
MVP area	0.34	0.47	0.35	0.48	-0.01
Polyg. household	0.31	0.46	0.32	0.47	-0.02
Monog. household	0.59	0.49	0.57	0.50	0.02
One-parent household	0.10	0.30	0.10	0.30	-0.00
Years school of HH	1.24	3.14	1.12	3.21	0.12
HH female	0.09	0.28	0.10	0.29	-0.01
HH age	46.97	12.93	48.54	12.48	-1.57**
N of hh members	8.18	3.08	8.96	4.05	-0.78***
N of children hh members	4.64	2.12	5.13	2.78	-0.49***
Ratio female members over total hh members	0.51	0.16	0.49	0.16	0.02**
Ratio female children over total children	0.48	0.26	0.47	0.26	0.02
N of households	1,010		458		

Note: Values of household level characteristics based on the first round of the survey (2012). First sample is formed by those household with at least one child aged 5-18 undertaking one cognitive test. Second sample is formed by those households with eligible children in which none of the them took any cognitive test. ***p<0.01; **p<0.05; *p<0.1.

2.6 Empirical Strategy

This section introduces the empirical strategy for the estimation of the effect of cognitive skills on the allocation of school attendance across siblings using four rounds of data collected for the evaluation of the Millennium Villages Project in northern Ghana. The main challenge for the identification of this effect arises from the fact that the link between cognitive skills and school attendance is likely affected by reverse causality or unobservable factors such as parental preferences. Indeed, although the cognitive tests used to measure cognitive skills are designed with the intention of being unaffected by schooling, culture and family background (see for example [Raven et al. \(1994\)](#)), it is not possible to rule out the possibility that these measures are themselves affected by schooling or by any other differential treatment that some children within the household may receive ([Ayalew, 2005](#); [Glewwe, 1999](#); [Akresh et al., 2012](#)).

An additional challenge when estimating the effect of cognitive skills on years of schooling arises from the fact that approximately the 38% of the eligible children in the sample did not take any of the cognitive tests. The restriction of the sample used in the analysis to only the children that took the cognitive tests could lead to a problem of sample selection bias. In line with this hypothesis, table 2.1 shows that those children that took at least one cognitive test are different from those that did not in terms of key characteristics such as labour force participation, sibling rank or school attendance.

To overcome these challenges, I follow [Cueto et al. \(2014\)](#) and propose the two steps procedure developed in [Heckman \(1979\)](#). In the first stage, I estimate equation 2.3 using a Probit model:

$$TakeCogTest_{i,h} = \omega_1 X_{i,h} + \omega_2 Z_h + u_{i,h} \quad (2.3)$$

where $TakeCogTest_{i,h}$ is a dummy variable that is equal to 1 if child i in household h took the cognitive test and X is a vector of child-level variables that are likely to affect the probability of taking the cognitive tests. This vector of variables includes school attendance, labour force participation, years of schooling, gender of the child, a proxy for health investments, sibling rank, labour force participation and a vector of dummies indicating the age in completed years of the child. Z_h is a vector of household level characteristics that could affect the probability of taking the cognitive tests. The vector of

household level variables includes per capita expenditure, an index of exposure to economic shocks, polygamy status, education and gender of the household head, household size, the number of children in the household, the number of male household members under 18 years and enumerator, village and time at survey fixed effects⁹. Both the individual and the household level characteristics included in the regression as right-hand side variables are measured in the first round of the survey (2012).

Once equation 2.3 is estimated, the next step is the calculation of the Inverse Mills Ratio (IMR), also known as selectivity correction term. The latter is calculated for every child as the coefficient of the standard normal density function of the predicted probability of taking cognitive test for every child divided by the standard normal cumulative distribution function of the predicted probability of taking the cognitive test for the same child.

In the second stage, I estimate using OLS the following equation:

$$YearsSchool_{i,h} = \omega_1 CognitiveSkills_{i,h} + \omega_2 X_{i,h} + \delta_h + u_{i,h} \quad (2.4)$$

where $YearsSchool_{i,h}$ is the number of years that the child i in household h went to school between the years 2013 and 2015, $CognitiveSkills_{i,h}$ indicates the age-adjusted z-score in the cognitive test taken in 2012 for child i and $X_{i,h}$ is a vector of child level characteristics in 2012 that includes gender, a proxy for health investments, sibling rank, a set of dummy variables for age, the IMR, years of schooling and school attendance. The specification also includes a dummy for every specific household surveyed. This vector of household fixed effects dummies account for differences in all the factors that are constant within a household. $u_{i,h}$ is the error term in the regression. The parameter of interest is ω_1 . Since household fixed effects are included in the regression, ω_1 yields the average effect on the number of years that the child attended school between 2013 and 2015 of an increase in one standard deviation in the cognitive test score of the child relative to the mean score of his siblings. Note that since the values of the variable $CognitiveSkills$ (measured in 2012) are not affected by school attendance between 2013 and 2015 and because the specification is already accounting for years of education and school attendance in 2012, the parameter

⁹The term fixed effects in this context implies that a dummy variable is included in the specification for every specific value found in the data for the variable of interest. For example, village fixed effects implies that a dummy is included in the specification for every village surveyed. The time at survey fixed effects is a vector of dummies for the hour of the day in which the enumerators started surveying the household.

ω_1 is not affected by the reverse causality problem. Furthermore, sample selection bias in equation 2.4 is addressed through the inclusion of the Inverse Mills Ratio (Heckman, 1979) as a control variable.

A potential limitation of the identification strategy presented in equation 2.4 is that in addition to innate cognitive ability and school investments, the cognitive skills measured in 2012 might be affected by other human capital investments. The household fixed effects may account for this problem if the investments are equal across siblings. As described in Akresh et al. (2012), the problem would arise if for example parents allocate more health investments to those siblings that have worse cognitive skills. If this is the case and the specification does not account adequately for health investments, the parameter that measures the effect of cognitive skills on schooling would be biased downwards. The best way I can cope with this problem is through including in the regression a dummy variable that takes the value of 1 if the child is registered with the National Health Insurance Scheme (NHIS) and 0 otherwise. The NHIS is a health insurance provided by the state. Registration with the NHIS is at the individual level and it is subject to a yearly fee that varies depending on the socioeconomic situation of the household. In this context, I use this variable to account for within household variation in the health investments received by the children.

2.7 Results

Table 2.3 presents the results of the first stage regression. The table shows that the child-level determinants of cognitive tests uptake are the same for the three tests examined. Being male, school attendance and lower age are factors positively associated with the uptake of cognitive tests. In the specifications with household fixed effects, the health insurance variable is also positive and statistically significant. Among the household level determinants, per capita expenditure, the number of male household members aged below 18 and whether the household head is male for the DSB test are negatively associated with the probability of children undertaking the cognitive tests.

Table 2.3: First stage (Probit coefficients): Participation of children in cognitive skills tests (children 5-18 in 2012).

	(1) Take Raven Test	(2) Take Raven Test	(3) Take DSF test	(4) Take DSF test	(5) Take DSB test	(6) Take DSB test
Female	-0.130*** (0.049)	-0.151** (0.076)	-0.127*** (0.049)	-0.139* (0.075)	-0.117** (0.048)	-0.141* (0.074)
Years of school 2012	-0.014 (0.011)	-0.010 (0.019)	-0.016 (0.011)	-0.017 (0.019)	-0.017 (0.011)	-0.022 (0.019)
Attendance 2012	0.438*** (0.070)	0.570*** (0.126)	0.411*** (0.071)	0.564*** (0.128)	0.426*** (0.070)	0.543*** (0.126)
Health insurance scheme	0.085 (0.064)	0.506*** (0.142)	0.051 (0.065)	0.501*** (0.146)	0.033 (0.064)	0.450*** (0.137)
Paid work	0.005 (0.068)	-0.102 (0.129)	-0.074 (0.067)	-0.141 (0.123)	-0.046 (0.067)	-0.131 (0.124)
Sibling rank	-0.011 (0.013)	-0.007 (0.020)	-0.021* (0.013)	-0.017 (0.020)	-0.011 (0.013)	-0.005 (0.020)
Age	-0.040*** (0.009)	-0.053*** (0.015)	-0.040*** (0.009)	-0.054*** (0.015)	-0.040*** (0.009)	-0.050*** (0.015)
Ln P/c expenditure	-0.103** (0.047)		-0.129*** (0.049)		-0.122** (0.049)	
Economic shock index (0-5)	0.019 (0.027)		0.039 (0.026)		0.025 (0.027)	
Polyg. household	0.022 (0.066)		0.036 (0.070)		-0.005 (0.068)	
N of children hh members	-0.010 (0.025)		-0.031 (0.025)		-0.018 (0.025)	
N of children-boys hh members	-0.040* (0.024)		-0.049** (0.024)		-0.046* (0.024)	
N of hh members	-0.017 (0.017)		-0.001 (0.017)		0.001 (0.017)	
Years school of HH	0.000 (0.008)		-0.001 (0.008)		0.003 (0.008)	
HH female	0.171 (0.106)		0.147 (0.104)		0.194* (0.106)	
HH age	-0.000 (0.002)		0.000 (0.002)		-0.000 (0.002)	
Enumerator FE	Yes	No	Yes	No	Yes	No
Time at survey FE	Yes	No	Yes	No	Yes	No
Household FE	No	Yes	No	Yes	No	Yes
Village FE	Yes	No	Yes	No	Yes	No
Observations	3,939	2,655	3,931	2,605	3,946	2,663

Note: Standard errors clustered at the household level are reported in parentheses. ***p<0.01; **p<0.05; *p<0.1.

Table 2.4 reports the estimates for equation 2.4 using the three different measures of cognitive skills collected in 2012: the score in Raven, DSF and DSB tests. The dependent variable in the analysis is the number of years that a child attended school between 2013 and 2015, calculated as the difference in years of schooling reported in 2015 and in 2012. Columns 1 to 3 of the table show the results for the three cognitive skills tests when equation 2.4 is estimated without accounting for non-random uptake of the cognitive tests. The results suggest that the effect of cognitive skills is positive and statistically significant at the 1%: an increase of one standard deviation in the score in cognitive tests relative to

the average score of the rest of the siblings in the household is associated with 0.126-0.178 more years of schooling between 2013 and 2015, depending on the cognitive measure used. When the IMR is included as a control variable to account for sample selection bias, the estimates displayed in columns 4 to 6 are very similar, with the effect of one standard deviation increase in the score of the cognitive tests relative to the average score of the rest of the siblings ranging between 0.128-0.178 years of schooling. Interestingly, the IMR is not statistically significant in these equations suggesting that at least in these estimates, the observed non-random uptake of the cognitive tests does not bias the estimation of the effect of cognitive skills on the allocation of schooling across siblings.

Table 2.4: Child cognitive skills and years of schooling 2012-2015 (children 5-18 in 2012).

	(1) Years of school 2015-2012	(2) Years of school 2015-2012	(3) Years of school 2015-2012	(4) Years of school 2015-2012	(5) Years of school 2015-2012	(6) Years of school 2015-2012
Female	0.054 (0.070)	0.115 (0.072)	0.087 (0.071)	0.010 (0.121)	0.071 (0.114)	0.079 (0.113)
Attendance 2012	0.667*** (0.142)	0.657*** (0.147)	0.665*** (0.148)	0.813** (0.403)	0.823** (0.374)	0.717* (0.414)
Years of school 2012	-0.114*** (0.025)	-0.119*** (0.025)	-0.123*** (0.024)	-0.119*** (0.027)	-0.128*** (0.028)	-0.125*** (0.029)
Health insurance scheme	0.380*** (0.141)	0.348** (0.148)	0.409*** (0.139)	0.429*** (0.158)	0.389*** (0.150)	0.415*** (0.141)
Sibling rank	0.002 (0.025)	-0.003 (0.024)	-0.004 (0.025)	-0.000 (0.026)	-0.009 (0.028)	-0.005 (0.026)
Raven age-adjusted z-score	0.126*** (0.041)			0.128*** (0.041)		
DSF age-adjusted z-score		0.167*** (0.046)			0.176*** (0.046)	
DSB age-adjusted z-score			0.178*** (0.042)			0.178*** (0.043)
IMR				0.559 (1.592)	0.639 (1.427)	0.201 (1.540)
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,147	2,061	2,073	2,125	2,034	2,058
R-squared	0.576	0.573	0.580	0.576	0.575	0.579

Note: Standard errors clustered at the household level are reported in parentheses.***p<0.01;**p<0.05;*p<0.1.

A potential critique to the analysis reported in table 2.4 is that the dependent variable measures school progression, which may not only be driven by school attendance but also by the cognitive skills of the child, which may directly influence his school marks and promotion to the next school year. Although grade promotion is automatic until senior secondary school and only less than 2% of the sample is currently or ever enrolled in post-junior secondary school educational levels¹⁰, it might be possible that the effects found are

¹⁰The educational system in Ghana is discussed in section 2.4.

driven by the most able children passing the BECE exam and accessing senior secondary school.

Table 2.5: Child cognitive skills and years of schooling 2012-2015 (Children 5-12 in 2012).

	(1) Years of school 2015-2012	(2) Years of school 2015-2012	(3) Years of school 2015-2012
Female	0.048 (0.204)	0.102 (0.181)	0.097 (0.182)
Attendance 2012	0.656 (0.715)	0.587 (0.581)	0.567 (0.668)
Years of school 2012	-0.154*** (0.044)	-0.153*** (0.047)	-0.153*** (0.046)
Sibling rank	0.025 (0.038)	0.028 (0.043)	0.017 (0.036)
Health insurance scheme	0.368* (0.217)	0.281 (0.183)	0.354** (0.167)
IMR	-0.122 (2.971)	-0.266 (2.351)	-0.684 (2.623)
Raven age-adjusted z-score	0.116** (0.055)		
DSF age-adjusted z-score		0.174*** (0.064)	
DSB age-adjusted z-score			0.211*** (0.065)
Age fixed effects	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes
Observations	1,342	1,284	1,304
R-squared	0.630	0.634	0.635

Note: Standard errors clustered at the household level are reported in parentheses. ***p<0.01; **p<0.05; *p<0.1.

To explore this possibility, I conduct the following two analyses. First, I re-estimate equation 2.4 using only the sample of children aged 12 or below in 2012. During the period studied, these children were only eligible for pre-school, primary or junior secondary education, where there is automatic course promotion at the end of the school course. The results of this analysis are reported in table 2.5. The coefficients of the cognitive tests are positive, statistically significant at conventional confidence levels and similar in terms of magnitude to those reported for the whole sample in table 2.4, suggesting that the effect of cognitive skills on years of schooling during the period 2013-2015 identified in table 2.4 is not driven by less able children being unable to achieve grade promotion at the end of

the school year. Second, I re-estimate equation 2.4 using as dependent variable the sum between 2013, 2014 and 2015 of a dummy variable collected yearly that is equal to 1 if the child attended school at least once during the last 12 months. The estimates for this analysis are reported in table 2.9 in appendix 2.A and show consistent results. The effect of an increase in one standard deviation in the score obtained in cognitive tests relative to the rest of the siblings in the household is positive and overall, statistically significant at 10%. The magnitude of this effect ranges between 0.046 and 0.060 additional years of school attendance during the period 2013-2015, depending on the cognitive test used and on whether the selectivity correction term is included as an additional control variable.

The results presented in this section are consistent with the main prediction of Becker's model: parents reinforce cognitive differences with their human capital investments. In a rural context where households have more than one child in school age and do not send all of them to school, parents are *ceteris paribus* more likely to concentrate school investments in the most able children in the household. This finding is also consistent with the results of previous studies that assess empirically the role of cognitive skills in the intra-household allocation of educational investments in Burkina Faso (Akresh et al., 2012), Ethiopia (Ayalew, 2005) and the US (Kim, 2005).

2.8 Heterogeneous Effects

This section examines whether the effect of cognitive skills on the allocation of schooling across siblings in a household depends on the gender of the child or on different household level characteristics. First, I test whether in a context with larger rates of school attendance among girls than among boys, the cognitive skills reinforcing mechanism operates differently across gender. Second, I investigate whether the strength of this reinforcing mechanism varies across polygynous and non-polygynous, larger and smaller and wealthier and poorer households.

2.8.1 Gender

In the sample, school attendance is more widespread and the average number of years of schooling is larger for girls (84% and 3.9) than for boys (77% and 3.5). GSS (2015) shows that the larger rates of attendance to primary school among girls than among boys observed

in the data is a well-established pattern in the Northern Region and in the Upper-East Region, the areas where the districts sampled are located. At the national level, Ghana has no gender differences in terms of net enrolment rates in primary education. One possible explanation for this could be the launch in the late 1990's and early 2000's of The Ghana Education Trust Fund (GETFUND) and the Compulsory Universal Basic Education (FCUBE) programme promoting school attendance in Ghana, particularly among girls and children from poor households (GSS, 2015).

The conceptual framework described in section 2.2 would predict a differential effect of cognitive skills on school investments in boys and in girls if the degree in which the returns to school for the child or for the parents depend on cognitive skills is different for boys and girls. The latter premise would be consistent with low levels of labour force participation for women. If returns to school depend more strongly on cognitive skills for boys, the model would predict that the allocation of investments in education across siblings reinforces more strongly cognitive differences between boys than between girls.

At this point, it is important to remark that a stronger effect of cognitive skills on school investments for boys would not imply that returns to education are lower for women than for men or that households find more profitable to allocate educational investments to boys than to girls. Rather, I am testing whether cognitive skills play a different role in understanding the intra-household allocation of schooling depending on the gender of the sibling. Thus, a weaker effect of cognitive skills on school investments in women would not be incompatible with the larger rate of school attendance for girls observed in the sample, particularly when considering that the economic returns to education for girls and their parents are not only limited to the labour market¹¹.

To test whether the effect of cognitive skills on the allocation of school investments across siblings is different for boys and girls, I propose two different tests. First, I estimate the following specification:

$$YearsSchool_{i,h} = \omega_1 CognitiveSkills_{i,h} + \omega_2 CognitiveSkills_{i,h} * Female_{i,h} + \omega_3 X_{i,h} + \delta_h + u_{i,h} \quad (2.5)$$

where $CognitiveSkills_{i,h} * Female_{i,h}$ is an interaction term of the score in the cognitive

¹¹For example, using data from Zimbabwe, Ashraf et al. (2016) show that education has large and positive effect in the bride price in the marriage market.

tests with a dummy variable that is equal to 1 if the child is a girl. X is a vector of child level variables that includes among others, a dummy variable that indicates the gender of the child and the IMR. In equation 2.5, the parameter ω_1 indicates the effect on school attendance for boys of increasing in one standard deviation the score in the cognitive test relative to the rest of the siblings. The effect for girls is yielded by the sum of the parameters $\omega_1 + \omega_2$. The parameter of interest in the equation would be ω_2 . If the latter parameter is larger (lower) than 0, the effect of cognitive skills on years of schooling would be larger (lower) for girls than for boys.

Table 2.6: Effect by gender: child cognitive skills and years of schooling 2012-2015 (children 5-18 in 2012).

	(1) Years of school 2012-2015	(2) Years of school 2012-2015	(3) Years of school 2012-2015	(4) Years of school 2012-2015	(5) Years of school 2012-2015	(6) Years of school 2012-2015
Female	0.007 (0.121)	0.066 (0.115)	0.081 (0.113)			
Attendance 2012	0.821** (0.402)	0.833** (0.376)	0.721* (0.415)	1.511* (0.892)	1.559* (0.888)	1.512 (0.968)
Years of school 2012	-0.122*** (0.027)	-0.128*** (0.028)	-0.126*** (0.029)	-0.062 (0.065)	-0.114 (0.076)	-0.103 (0.084)
Health insurance scheme	0.428*** (0.156)	0.390*** (0.151)	0.410*** (0.142)	-0.244 (0.412)	-0.409 (0.462)	-0.310 (0.462)
Sibling rank	-0.002 (0.026)	-0.010 (0.028)	-0.006 (0.026)	0.016 (0.124)	-0.008 (0.127)	0.054 (0.125)
IMR	0.633 (1.590)	0.712 (1.437)	0.216 (1.546)	3.886 (3.445)	4.325 (3.132)	3.627 (3.439)
Raven age-adjusted z-score	0.201*** (0.052)			0.230* (0.137)		
Female \times Raven	-0.155** (0.069)			-0.042 (0.196)		
DSF age-adjusted z-score		0.210*** (0.059)			0.026 (0.259)	
Female \times DSF		-0.078 (0.074)			0.257 (0.340)	
DSB age-adjusted z-score			0.209*** (0.056)			0.162 (0.188)
Female \times DSB			-0.067 (0.074)			0.086 (0.363)
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,125	2,034	2,058	252	249	234
R-squared	0.579	0.575	0.579	0.723	0.693	0.712

Note: Standard errors clustered at the household level are reported in parentheses.***p<0.01;**p<0.05;*p<0.1.

Second, I estimate equation 2.5 restricting the analysis to those households with all the household individuals aged 5-18 years either all girls or all boys. In this estimation, the variable $CognitiveSkills_{i,h} * Female_h$ is an interaction term of cognitive skills with a dummy variable defined at the household level that is equal to 1 if all the children aged 5-18 in the household are girls and 0 if all the children 5-18 are boys. In this case, ω_1

indicates the effect on school attendance of increasing in one standard deviation the score in the cognitive test relative to the rest of the siblings in households where all the children are boys. The effect in households where all the children are girls is yielded by the sum of the parameters $\omega_1 + \omega_2$.

The estimates for these two analyses are reported in table 2.6 and show mixed results. On the one hand, the interaction term *Female* \times *CognitiveSkills* in columns 1 to 3 is consistently negative suggesting that the effect of cognitive skills on the allocation of schooling across siblings is larger for boys than for girls although this coefficient is only statistically significant at conventional confidence levels for the Raven test. On the other hand, the estimations reported in columns 4 to 6 yield mixed results. While the interaction term is negative for the Raven test, it is positive for the digit span tests although in none of the cases is statistically significant at conventional confidence levels. In any case, the sample size used in this second analysis is small and thus, the results should not be taken as conclusive.

2.8.2 Household Consumption

An additional prediction of Becker's model is that the reinforcing pattern in the allocation of human capital investments would be weaker (if any) in poor households because nonhuman capital resources are limited in these households and therefore, disadvantaged households would not be able to compensate less able children with larger nonhuman capital transfers. At this point, the reinforcing mechanism in households that cannot make nonhuman capital investments would depend on the degree of parents' aversion towards sibling inequality: the stronger the aversion for sibling inequality, the smaller the difference in schooling investments received by more and less able siblings, and the larger the differential effect of cognitive skills on the allocation of schooling across siblings in poorer and wealthier families. To test this hypothesis, I estimate the following specification:

$$YearsSchool_{i,h} = \omega_1 CognitiveSkills_{i,h} + \omega_2 CognitiveSkills_{i,h} * Consum_{i,h} + \omega_3 X_{i,h} + \delta_h + u_{i,h} \quad (2.6)$$

where *CognitiveSkills*_{*i,h*} \times *Consum*_{*h*} is an interaction term of cognitive skills with a set of dummy variables defined at the household level that capture whether the household is in

the first, second or third tercile in the distribution of per capita household consumption. In equation 2.6, the parameter ω_1 yields the effect of cognitive skills on the allocation of schooling across siblings in households in the top tercile of the per capita consumption distribution. The vector of parameters ω_2 are the estimates of first importance in this analysis. If the latter parameters are larger (lower) than 0, the effect of cognitive skills on schooling in these households would be larger (lower) than in the households in the richest tercile.

Table 2.7: Effect by household characteristics: child cognitive skills and years of schooling 2012-2015 (children 5-18 in 2012).

	(1) Years of school 2012-2015	(2) Years of school 2012-2015	(3) Years of school 2012-2015
Female	0.104 (0.115)	0.108 (0.115)	0.113 (0.115)
Attendance 2012	0.491 (0.377)	0.492 (0.378)	0.476 (0.378)
Years of school 2012	-0.108*** (0.027)	-0.108*** (0.027)	-0.108*** (0.027)
Sibling rank	0.003 (0.027)	0.004 (0.027)	0.004 (0.027)
IMR	-0.880 (1.465)	-0.902 (1.471)	-0.946 (1.470)
Raven age-adjusted z-score	0.096* (0.054)	0.139* (0.077)	0.203** (0.092)
Polygamous hh \times Raven score	0.082 (0.083)		
HH<4 children \times Raven score		-0.021 (0.118)	
HH 4-6 children \times Raven score		-0.012 (0.096)	
P/c Consump(Terc. 1) \times Raven score			-0.114 (0.110)
P/c Consump(Terc. 2) \times Raven score			-0.071 (0.112)
Age fixed effects	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes
Observations	2,125	2,125	2,125
R-squared	0.573	0.573	0.573

Note: Standard errors clustered at the household level are reported in parentheses.***p<0.01;**p<0.05;*p<0.1.

The results of this analysis are reported in column 3 of table 2.7. Overall, they are

consistent with the prediction of Becker's model under some degree of aversion for sibling inequality. The effect of cognitive skills on school investments in the richest tercile of households is statistically significant at the 5%. The parameters that measure the differential effect in households in the first and second tercile are negative indicating that on average, the lower the per capita consumption in the household, the smaller the effect of cognitive skills on the allocation of schooling across siblings. However, although they have the expected negative sign, the coefficients that measure the differential effects across terciles are not statistically significant at conventional significance levels.

2.8.3 Polygyny

Defined as the marital life of one men and more than one women, the practice of polygyny is widespread in West Africa. Although polygamy is not as high in Ghana as it is in other West African countries such as Guinea Conakry or Senegal, the 31% of the households in the data used in this study are polygynous

One assumption behind the predictions of Becker's model is the alignment of incentives across the parents of the children. However, polygyny could break the alignment of incentives across parents if every mother prefer her own children to be more educated than the children of the other women in the household regardless of their cognitive skills. To examine whether the misalignment of incentives across parents in polygynous household affects the role of cognitive skills in the allocation of schooling investments across siblings, I estimate the following equation:

$$YearsSchool_{i,h} = \omega_1 CognitiveSkills_{i,h} + \omega_2 CognitiveSkills_{i,h} * Polyg_h + \omega_3 X_{i,h} + \delta_h + u_{i,h} \quad (2.7)$$

where $CognitiveSkills_{i,h} * Polyg_h$ is an interaction term of the score in the Raven test with a dummy variable defined at the household level that is equal to 1 if the child lives in a polygynous household and 0 otherwise. In equation 2.7, the parameter ω_1 indicates the effect on school attendance of increasing in one standard deviation the cognitive test score in non-polygynous households. The effect in polygynous households is yielded by the sum of the parameters $\omega_1 + \omega_2$. The key parameter is ω_2 . If this parameter is larger (lower) than 0, the effect of cognitive skills on school attendance in polygynous households would

be larger (lower) than in non-polygynous households.

Column 1 of table 2.7 shows the estimates for equation 2.7. The coefficient for ω_2 is positive suggesting that cognitive ability relative to siblings is a stronger determinant of educational investments received in polygynous household. However, although the magnitude of the coefficient is non-negligible, it is also statistically indistinguishable from 0 at conventional confidence levels. The results suggest that the misalignment of incentives across parents arguably existent in polygynous households does not seem to weaken the role of cognitive skills in the allocation of schooling across siblings. Two potential explanations for this could be that more cognitively able wives (that may have children with better innate cognitive skills) achieve higher educational investments for their children or that fathers in polygynous households decide alone on the allocation of schooling investments across siblings.

Although the survey collects information on the nature of the relation of every household member with the household head, the survey does not collect sufficient information to identify confidently the mother of most of the children in polygynous households. This fact hinders the possibility of providing a deeper insight into the role played by mother characteristics or wife rank in the intra-household allocation of school investments in polygynous households.

2.8.4 Household Size

Children with a lot of siblings in school age may benefit from economies of scale in school costs (e.g. books, uniforms) or in household chores. On the other hand, more siblings can increase competition for scarce resources. In order to determine whether the effect of cognitive skills on the allocation of schooling investments across siblings depends on the number of children in the household, I estimate equation 2.8:

$$YearsSchool_{i,h} = \omega_1 CognitiveSkills_{i,h} + \omega_2 CognitiveSkills_{i,h} * HhSize_h + \omega_3 X_{i,h} + \delta_h + u_{i,h} \quad (2.8)$$

where $CognitiveSkills_{i,h} * HhSize_h$ is a vector of interaction terms of the score in the Raven test with a set of dummy variables defined at the household level that indicate

whether less than 4 children¹², between 4 and 6 children or more than 6 children live in household h . The thresholds used to define the three household size dummies resulted in a division of the sample of households in three equal parts. In equation 2.8, the parameter ω_1 indicates the effect of cognitive skills on the allocation of school attendance across siblings in the reference group, that is formed by those households with more than 6 children. On the other hand, the vector of parameters ω_2 measures the differential effect of cognitive skills on school attendance in households with 4 to 6 children and in households with less than 4 children relative to the reference group. If these parameters are larger (lower) than 0, the effect of cognitive skills on school attendance in these groups of households would be larger (lower) than in the reference group.

The results of this analysis are reported in column 2 of table 2.7. The coefficients show that on average, the larger the household size, the larger the effect of Raven score on years of schooling. For example, the estimated effect of an increase in one standard deviation in the Raven test score on years of schooling between 2012 and 2015 is 0.139 in households with more than 6 children. This coefficient is statistically significant at the 10%. The estimates provided show that the effect is 0.127 (0.139-0.012) in households with 4 to 6 children and 0.118 (0.139-0.021) in households with less than 6 children. However, it is important to remark that none of the coefficients that measure the differential effect of cognitive skills on the allocation of schooling in smaller and larger households is statistically significant at conventional confidence levels. Therefore, in the light of the results, it is not possible to reject the hypothesis that the magnitude of the effect of cognitive skills on the allocation of schooling across siblings does not depend on household size.

2.9 Conclusions

This study examines empirically a prediction of the model of intra-household allocation of resources developed in the seminal paper [Becker \(1981\)](#): with their investment in schooling, parents reinforce cognitive differences between siblings. Relying on 4 rounds of a household panel survey conducted in 103 villages of the north of Ghana, I find evidence that cognitive skills strongly determine the allocation of schooling across siblings. In the preferred set of specifications, an increase of one standard deviation in the score of cognitive tests relative to the rest of the siblings in the household, raises the number of years of schooling

¹²A child is defined as a household member aged below 18 years.

attended in the following three years by 0.128-0.178, depending on the cognitive measure used. These results are consistent with the main prediction of Becker's model and are in line with the results found in previous studies.

The evidence suggests that policies aiming to increase school attendance in northern Ghana should take into account that parents target their educational investments towards the most capable children rather than spreading these investments equally among all the children in the household. In this context, demand-side educational interventions such as conditional cash transfers that promote school attendance among less able siblings would probably be more effective in promoting universal schooling than supply side interventions such as reductions in class size ([Akresh et al., 2012](#)).

The study also investigates whether the magnitude of the effect of cognitive skills on the allocation of schooling across siblings depends on the gender of the child or on different household characteristics. On the one hand, the analysis suggests that the effect of cognitive skills on schooling is larger for boys although the results should only be interpreted as suggestive and they are not unambiguous. On the other hand, I do not find any evidence that the effect of cognitive skills on the allocation of schooling across siblings is significantly different in richer and poorer, polygynous and non-polygynous, and larger and smaller households.

Appendix 2.A Appendix

Table 2.8: First stage (Probit coefficients): Participation of children 5-18 in 2012 in cognitive skills tests (school attendance).

	(1) Take Raven Test	(2) Take Raven Test	(3) Take DSF test	(4) Take DSF test	(5) Take DSB test	(6) Take DSB test
Female	-0.133*** (0.049)	-0.150* (0.077)	-0.138*** (0.049)	-0.154** (0.076)	-0.124** (0.048)	-0.150** (0.076)
Years of school 2012	-0.009 (0.012)	-0.003 (0.020)	-0.012 (0.012)	-0.009 (0.020)	-0.012 (0.011)	-0.015 (0.020)
Attendance 2012	0.400*** (0.071)	0.539*** (0.129)	0.378*** (0.073)	0.529*** (0.130)	0.394*** (0.071)	0.509*** (0.127)
Health insurance scheme	0.095 (0.064)	0.490*** (0.143)	0.054 (0.064)	0.468*** (0.146)	0.032 (0.063)	0.411*** (0.139)
Paid work	-0.009 (0.069)	-0.088 (0.132)	-0.082 (0.068)	-0.131 (0.125)	-0.054 (0.068)	-0.118 (0.126)
Sibling rank	-0.012 (0.013)	-0.005 (0.020)	-0.022* (0.013)	-0.014 (0.020)	-0.011 (0.013)	-0.001 (0.020)
Age	-0.042*** (0.009)	-0.057*** (0.015)	-0.042*** (0.009)	-0.056*** (0.015)	-0.041*** (0.009)	-0.052*** (0.015)
Ln P/c expenditure	-0.100** (0.047)		-0.119** (0.049)		-0.118** (0.049)	
Economic shock index (0-5)	0.012 (0.028)		0.034 (0.027)		0.020 (0.028)	
Polyg. household	0.027 (0.067)		0.044 (0.071)		0.005 (0.069)	
N of children hh members	-0.011 (0.025)		-0.030 (0.025)		-0.017 (0.025)	
N of children-boys hh members	-0.044* (0.025)		-0.056** (0.025)		-0.051** (0.024)	
N of hh members	-0.015 (0.017)		-0.001 (0.017)		0.001 (0.017)	
Years school of HH	0.001 (0.008)		0.000 (0.008)		0.004 (0.008)	
HH age	-0.000 (0.002)		0.000 (0.002)		-0.000 (0.002)	
HH female	0.206* (0.108)		0.162 (0.106)		0.224** (0.109)	
Enumerator FE	Yes	No	Yes	No	Yes	No
Time at survey FE	Yes	No	Yes	No	Yes	No
Household FE	No	Yes	No	Yes	No	Yes
Village FE	Yes	No	Yes	No	Yes	No
Observations	3,875	2,584	3,867	2,528	3,882	2,588

Note: Standard errors clustered at the household level are reported in parentheses. ***p<0.01; **p<0.05; *p<0.1.

Table 2.9: Child cognitive skills and school attendance 2013-15 (children 5-18).

	(1)	(2)	(3)	(4)	(5)	(6)
	N years attended school 2013-2015	N years attended school 2013-2015	N years attended school 2013-2015	N years attended school 2013-2015	N years attended school 2013-2015	N years attended school 2013-2015
Female	0.042 (0.044)	0.062 (0.045)	0.051 (0.043)	0.067 (0.091)	0.111 (0.079)	0.151** (0.073)
Attendance 2012	1.236*** (0.121)	1.271*** (0.125)	1.280*** (0.121)	1.147*** (0.289)	1.105*** (0.250)	0.941*** (0.261)
Years of school 2012	0.047*** (0.016)	0.037** (0.017)	0.040** (0.016)	0.049*** (0.017)	0.041** (0.019)	0.049*** (0.017)
Sibling rank	0.020 (0.016)	0.021 (0.017)	0.024 (0.016)	0.022 (0.018)	0.028 (0.020)	0.030* (0.016)
Health insurance scheme	0.385*** (0.106)	0.387*** (0.108)	0.378*** (0.108)	0.361*** (0.129)	0.364*** (0.114)	0.341*** (0.112)
Raven age-adjusted z-score	0.046* (0.025)			0.047* (0.025)		
DSF age-adjusted z-score		0.052* (0.031)			0.050 (0.031)	
DSB age-adjusted z-score			0.060** (0.028)			0.056* (0.029)
IMR				-0.443 (1.253)	-0.789 (0.928)	-1.538 (0.960)
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,132	2,046	2,057	2,110	2,019	2,042
R-squared	0.745	0.750	0.758	0.745	0.751	0.759

Note: Standard errors clustered at the household level are reported in parentheses.***p<0.01;**p<0.05;*p<0.1.

Chapter 3

The Lasting Effects of Natural Disasters on Property Crime: Evidence from the 2010 Chilean Earthquake

3.1 Motivation

Aside from the images of destruction, one aspect often displayed by televisions and newspapers in the aftermaths of natural disasters are scenes of chaos and looting. The evidence suggests that the disorders that followed natural disasters such as the Katrina hurricane or the 2010 Haiti earthquake seem to be closely linked to power cuts and to the collapse of the police, decreasing temporarily the cost of crime ([Frailing and Harper, 2007](#); [Friesema et al., 1979](#); [Kolbe et al., 2010](#)).

Although the *break in the social contract* often found in the aftermath of natural disasters is usually limited to a few days or even hours ([Quarantelli, 2001](#)), temporary reductions in the cost of crime could lead to lasting effects on crime rates if the personal cost of committing crime is a decreasing function of the number of previous crimes committed. Similarly, the negative effects of disasters on employment ([Belasen and Polachek, 2008](#)) could also decrease in the long-term the opportunity cost of crime. On the other hand, natural disasters may also strengthen community links ([Dynes and Quarantelli, 1980](#); [Bailey, 2009](#)), facilitating the adoption of community-based crime prevention strategies and increasing the cost of committing crime. With mixed effects on the benefits and costs of crime, the long-term impact of natural disasters on property crime is theoretically ambiguous and is therefore an empirical question.

I provide evidence on the lasting effects of natural disasters on property crime using as a case study the 8.8 Richter magnitude earthquake that struck the Centre-South of Chile in 2010. This earthquake caused 547 fatalities and economic damages estimated at USD 15-30 billions (UNEP, 2011). In the aftermath of the catastrophe, the most affected areas experienced looting episodes that involved hundreds of people and in response, the Chilean government deployed the army and declared a curfew in these municipalities. The main estimations presented in this study rely on difference in difference models comparing crime rates in municipalities close and far away from the epicentre and use pre- and post-earthquake data from 7 rounds of a household victimization survey conducted every year in 101 urban municipalities in Chile.

The results of the study reveal that, depending on the specification, exposure to a *very strong* earthquake intensity decreased the incidence of home burglary the year of the earthquake by 16%-30% relative to areas not directly affected by the disaster and that this effect remained constant over the 4 post-earthquake years studied. The results hold when other sources of crime data and other types of property crime are examined, ruling out the possibility of crime displacement from home burglary towards other types of property crime. The results are also robust to the use of alternative samples and earthquake intensity thresholds to define treatment and control municipalities.

I examine different mechanisms for lower property crime in earthquake affected areas. The results show that the main channel that drove the lasting contraction in property crime rates was the positive effect of the disaster on the strength of community life. The improvement in the social capital at the community level boosted the adoption of community-based measures to prevent crime, increasing the cost of crime in earthquake affected areas. Alternative mechanisms such as an increase in the number of policemen or a reduction in unemployment due to reconstruction activities in areas affected by the earthquake are rejected in the light of the results. Furthermore, the estimates also suggest that the lasting drop in the incidence of property crime was not caused by higher levels of incarceration as a consequence of the institutional efforts in the aftermath of the earthquake to address looting or by a rise in the perceived risk of crime boosting permanent adoption of crime prevention measures. Finally, the analysis shows that the persistent reduction in property crime rates in earthquake affected areas was not driven by a lasting effect of the deployment of the army or the curfew via a temporary increase in the cost of

crime with long-term term consequences.

This study is primarily related with the body of literature that investigates the effect of natural disasters on the incidence of crime. Indeed, the results are consistent with the informal guardianship theory developed in sociology that argues that natural disasters increase co-operation and the formation of social capital within damaged communities, increasing the provision of informal guardianship in these communities and therefore the cost of crime. In this context, the contribution of the study is twofold. First, unlike previous studies that examine the evolution of crime rates over a maximum post-disaster period of 12 months, this study explores the impact of natural disasters on property crime over a longer period of time (4 post-disaster years). Second, this is the first study that investigates empirically the mechanisms driving the effects of natural disasters on property crime, providing empirical evidence that supports the different hypotheses of the informal guardianship theory and showing the key role that social capital at the community level can play in the reduction of property crime.

The study is structured as follows. Section 3.2 introduces the conceptual framework. Section 3.3 discusses the existing evidence on the link between natural disasters and crime. Section 3.4 describes the context and the main political and social events that followed the 2010 Chilean earthquake. Section 3.5 presents the data and section 3.6 introduces the empirical strategy used to estimate the effect of the earthquake on property crime. Section 3.7 discusses the main results of the analysis and section 3.8 expands the analysis to other types of property crime using an alternative source of crime data. Section 3.9 explores the mechanisms through which the earthquake could have reduced property crime and section 3.10 concludes.

3.2 Conceptual Framework

Mainly developed in the field of sociology, there are two opposing streams of literature that set out different predictions about how crime rates evolve following natural disasters.

The first stream of authors hypothesises that crime rates increase following natural disasters. There are three main mechanisms through which this effect might operate. The first of them, known in the field of sociology as the *routine activities theory*, is described in Cohen and Felson (1979). They argue that natural disasters are followed by a rise in

crime rates because catastrophes increase the availability of suitable targets and reduce the presence of capable guardianship. Another mechanism that explain why crime rates could spike following natural disasters is that crime is more prevalent in those places characterized by the incapacity of the community to informally control crime due to factors such as residential instability that might be severely damaged by natural disasters (Zahran et al., 2009). This argument is interpreted in the context of the *social disorganization* theory developed in Shaw and McKay (1942). These two mechanisms can be embedded in traditional economics of crime models that describe crime rates as a function of crime's costs and benefits (Ehrlich, 1973; Becker, 1968): Through causing a temporary or permanent obstruction of law enforcement, generating power cuts and forcing some households to leave their dwellings, natural disasters decrease the cost of committing crime, leading to larger crime rates. The third path through which natural disaster may affect property crime is the labour market. If employment represents the opportunity cost of crime, the lasting negative effect of natural disasters on labour outcomes documented in Belasen and Polachek (2008) could boost the incidence of crime.

The second stream of the literature argues that crime rates do not raise and might even decrease following natural disasters. These authors highlight that although natural disasters may decrease the capacity of *formal* institutions such as the police to enforce the law, they also raise pro-social and altruistic behaviours (Quarantelli and Dynes, 1970), fostering co-operation and the formation of social capital within communities and increasing the level of informal guardianship (Cromwell et al., 1995). The authors argue that the rise in the level of informal guardianship offsets the potential harmful effects of natural disasters on crime arising from a reduced capacity of the police to enforce the law immediately after disasters or from the perverse effect of the disaster on other crime determinants, eventually leading to lower crime rates. In a traditional economics of crime theoretical model, the argument of these authors implies that far from reducing crime costs, the rise in the provision of informal guardianship in affected communities compensates the reduced capacity of formal institutions to provide capable guardianship, increasing the probability of apprehension and the cost of committing crime in these communities.

The two streams propose different channels through which natural disasters can affect the costs and benefits of crime with opposite directions. In the light of this literature, the effect of natural disasters on crime predicted in theoretical models of economics of crime

would be ambiguous, with the sign of the net effect depending on the superiority of some channels over others and highlighting that the effect of natural disasters on property crime is an empirical question.

3.3 Related Literature

The short-term evolution of crime rates following natural disasters has been empirically investigated in different studies, with mixed results.

Most of the studies addressing this question find that crime rates increase after natural disasters. For example, [Roy \(2010\)](#) exploits district-level panel data in India to investigate the incidence of violent and property crime in districts that experienced a natural disasters the same year. The paper shows that, overall, natural disasters are followed by increases in most types of property and violent crime. Using known to the police crime data, [Friesema et al. \(1979\)](#) show large increases in motor vehicle theft in Texas following hurricane Carla. [Frailing and Harper \(2007\)](#) find a spike in the incidence of burglary in New Orleans after hurricane Katrina, and [Kolbe et al. \(2010\)](#) suggest that the large earthquake that affected Haiti in 2010 triggered sexual assaults in the weeks following the disaster. [Leitner and Helbich \(2011\)](#) investigate the link between crime and natural disasters through studying the daily evolution of crime rates before, during and after two hurricanes that affected the city of Houston. The authors state that while burglary and motor-vehicle theft increased immediately before and after hurricane Rita, crime rates did not change before, during or after hurricane Katrina. They argue that the difference in effects might be driven by the fact that while an order of evacuation was issued before hurricane Rita, no such order was issued before or during hurricane Katrina in Houston. The empirical evidence supporting the suggestion that natural disasters are followed by an increase in crime rates is particularly strong for domestic and sexual violence offences such as child abuse ([Curtis et al., 2000](#)), sexual assault ([Kolbe et al., 2010](#)) or gender violence ([Peacock et al., 1997](#); [Enarson et al., 2006](#)).

However, the evidence is not homogeneous and there are some empirical studies that find either a decrease or a stagnation in crime rates after natural disasters. For example, using qualitative data collected one month after hurricane Andrew in Florida, [Cromwell et al. \(1995\)](#) show that although the hurricane increased the number of motivated offenders

and unprotected victims, it also boosted informal guardianship leading to sharp decreases in crime rates during the weeks that followed the hurricane. Similarly, [Siegel et al. \(1999\)](#) find that exposure to the 1994 Northridge earthquake in California did not increase the likelihood of suffering a violent or a property crime during the two months that followed the disaster. Although the evidence is mixed, most of the studies that explore the evolution of crime rates in New Orleans and neighbouring parishes after hurricane Katrina suggest that except for burglaries, property crime rates decreased the months following the disaster although the rates converged to pre-hurricane levels one year later ([Leitner et al., 2011](#); [Bailey, 2009](#))¹.

[Zahran et al. \(2009\)](#) bring the discussion a step forward arguing that the incidence of different types of crimes might evolve differently after natural disasters. Using county-level panel data from Florida and well-conducted fixed effects techniques, the paper provides evidence that while natural disasters tend to decrease property and violent crime the year of the disaster, they also raise the incidence of domestic violence.

Although the number of studies that explore the short-term effect of natural disasters on crime is large, most of these studies lack methodological rigour. For example, only two of the studies discussed ([Roy, 2010](#); [Zahran et al., 2009](#)) use a counterfactual approach to account for potential confounding factors and with one exception ([Kolbe et al., 2010](#)), the literature relies on crime data from police records. The use of police records could be problematic because changes in crime known to the police after natural disasters could be reflecting an effect of natural disasters on the probability of reporting crime to the police rather than on true crime rates. Furthermore, I am not aware of any previous study investigating whether the effects of natural disasters on crime expand over more than one year. Finally, and although some of the studies discuss them theoretically, this is the first study that explores empirically the mechanisms driving the effect of natural disasters on property crime.

3.4 The Context

The early morning of the 27th of February of 2010 an earthquake of 8.8 degrees in the Richter scale shook the Centre-South of Chile. The epicentre was located approximately

¹The evidence on crime dynamics after Katrina hurricane is mixed and some studies also show that crime rates one year after the disaster were larger than pre-hurricane rates, particularly for murder ([Van-Landingham, 2009](#)).

90 km north west of Concepción, the second largest Chilean city with a metropolitan population above 1,000,000 inhabitants. The earthquake was followed by a tsunami with waves striking approximately 500 km of the Chilean coast. Although the economic losses affected a total of 6 regions that included the 80% of the Chilean population, the regions of Biobio and Maule were particularly damaged by the earthquake and the tsunami (Larrañaga and Herrera, 2010b).

Different reports from the Chilean government, NGOs, universities and international organizations provide an estimation of the economic damages and human losses caused by the earthquake and the tsunami. Nahuelpan and Varas (2011) report that the earthquake and the tsunami that followed caused a total of 547 deaths. Contreras and Winckler (2013) attribute 181 of these deaths to the tsunami. Regarding the direct economic losses caused by the earthquake and the tsunami, UNEP (2011) estimates in USD 15-30 billions the damage caused to public and private assets, including 440,000 houses and numerous roads severely deteriorated (CEPAL, 2010). Although identifying the losses caused only by the tsunami is in most cases difficult, Contreras and Winckler (2013) argue that it damaged 17,392 houses in 24 different municipalities. The same report also highlights that the tsunami affected many coastal infrastructures including different harbours and piers and approximately 3,000 boats. Table 3.1 summarizes the main losses at the regional level for the six regions affected by these natural disasters.

The earthquake also caused water, power and telephonic cuts. Power cuts affected the 80% of the population and lasted between a few hours and three days in the most damaged areas of the country (OPM, 2010). After some looting episodes in the regions of Biobio and Maule, the 28th of February the Chilean government declared the state of emergency for 30 days in these two regions and a curfew in the municipalities that experienced looting episodes. Following the declaration, the army was deployed in urban areas of these regions, particularly in Biobio². Nonetheless, looting did not completely stop and pillage episodes were occasionally registered during the following week³. In total, there were looting events in 33 municipalities (Ormeño, 2010). Some of these episodes were documented by the media and involved hundreds of looters⁴.

²http://internacional.elpais.com/internacional/2010/02/28/actualidad/1267311602_850215.html

³<http://www.ambito.com/noticia.asp?id=510234>

⁴see for example http://www.24con.com/nota/37127_Saqueos_la_gente_se_lleva_desde_lechehastaplasmas

Qualitative data and media reports point out that the earthquake was followed by social chaos in heavily affected areas that ended with many people participating in looting events mainly towards big supermarkets and shops⁵. However, despite the limited capacity of the police to enforce the law the days following the earthquake, the looting of dwellings and habited places was a very rare event (Grandón et al., 2014; Larrañaga and Herrera, 2010b). Remarkably, these reports also document widespread pro-social and altruistic behaviours in the aftermath of the earthquake and communities organizing themselves to overcome earthquake catastrophic consequences.

Perhaps influenced by the media coverage of the post-disaster events, the 32% of the urban households interviewed for the 2010 ENUSC survey believed that the earthquake caused an increase in the incidence of crime at the national level during the same year. Interestingly, the percentage of households that reported such perception was higher in the areas far away from the earthquake epicentre (32% in control municipalities) than in areas close to it (27% in treatment municipalities).

Table 3.1: Fatalities and economic damage of the earthquake/tsunami by region

	Fatalities	% Dwellings severe damage	% Dwellings severe damage (I quintile)	% Dwellings severe damage (V quintile)	% HHs facing problem from earthquake/tsunami	% pop >18 with symptoms post-traumatic stress
Valparaíso	25	7.4	11.3	2.4	51.9	8.3
O'Higgins	53	12.2	12.5	7.5	67	22.3
Maule	280	20.7	26.3	12.8	92.9	21.4
Biobío	145	17.8	25.4	8.5	92.9	23.9
Araucanía	17	5.1	10.2	0.5	59.3	11.5
Metropolitana	27	4.8	6.5	3.0	56	6.5
All regions aff.	547	8.8	12.0	4.6	64.7	12.0

Source: Larrañaga and Herrera (2010a). Information on damages is only provided for the six regions affected by the earthquake. The regions of Tarapacá, Arica y Parinacota, Atacama, Coquimbo, Antofagasta, Los Ríos, Los Lagos, Aisén and Magallanes are not included in the survey because the authors concluded that they were not directly damaged by the earthquake or the tsunami. The I quintile indicates the quintile of the Chilean population with the lowest per capita income while the V quintile is the richest quintile of the Chilean population in terms of per capita income.

3.5 Data

The crime data used in the main analysis correspond to seven rounds of the Encuesta Nacional Urbana de Seguridad Ciudadana (ENUSC) survey for the period 2007-2013⁶. The ENUSC is a household survey conducted by the Chilean Ministry of Governance and

⁵ http://ciperchile.cl/2010/07/19/saqueadores_post_terremoto_ii_la_horda_que_nunca_llego_a_las_casas

⁶The first publicly available ENUSC survey was conducted in 2007.

applied every year to a cross section of more than 25,000 urban households living in the largest 101 Chilean municipalities. The survey collects household level information on victimization in the last 12 months for different types of crimes and on the adoption of individual and community-based measures to prevent crime.

The main advantage of the ENUSC data relative to crime data from police records is that while police records only include those offences reported to or unmasked by the police, the ENUSC survey captures both the crimes reported to and unreported to the police. On the other hand, the use of this dataset has two drawbacks. First, with the exception of home burglary, the exact location of each crime is not reported. This could be particularly problematic for the metropolitan areas of Santiago and Concepción where many individuals work and live in a different municipality. Second, the difference between some types of property crimes such as larcenies, burglaries or distraction theft is in many cases fuzzy. In consequence, some households might be unable to report reliably some specific types of crime to the enumerator. For these two reasons, I restrict the analysis conducted using the ENUSC database to home burglary; an offence that is unlikely to be confounded with other crimes by the households interviewed or the enumerator and for which the exact location is known.

Section 3.8 tests the robustness of the results and expands the analysis to other types of property crime including motor-vehicle theft, non-home burglary, larceny and robbery using crime data from police records. These records were obtained from the Subsecretaría de Prevención del Delito (SPD) in Chile and they report every month and year (a) the number of crimes known to the police in each of the 345 Chilean municipalities by type of crime⁷ and (b) the number of individuals apprehended by the police in every municipality.

The dataset on earthquake intensity is constructed using the geographical information provided by the Oficina Nacional de Emergencia del Ministerio del Interior y Seguridad Pública (ONEMI) on the coordinates, magnitude and depth of the earthquake hypocentre. The distance to the earthquake hypocentre is then used to predict the Modified Mercalli Intensity (MMI) at the municipality level using the method described in Barrientos (1980), that predicts earthquake intensity in a given place as a function of the distance from the place to the hypocentre and of the earthquake magnitude at its source⁸.

⁷The crime data are available at http://www.seguridadpublica.gov.cl/tasa_de_denuncias_y_detenciones.html

⁸Using data from 945 measurements of earthquake intensity in different places after 73 earthquakes $M_w > 5.5$ that struck Chile between 1906 and 1977, the paper estimates the following function that

In the main analysis, I define as treatment municipalities those exposed to a predicted $MMI \geq 7.5$. The expected damages associated with a $MMI = 7$ are negligible damage in buildings of good design and construction; slight to moderate in well-built ordinary structures and considerable damage in poorly build or badly designed structures⁹. However, I set the threshold in predicted $MMI \geq 7.5$ because the predicting method developed in Barrientos (1980) seems to overestimate intensities $MMI > 7$ for this particular earthquake (Astroza et al., 2010). Control municipalities are defined as those exposed to a predicted $MMI < 5.75$. I set this threshold to define control municipalities because the damages associated with a $MMI < 6$ are minimum (Astroza et al., 2010) and Mercalli intensities are usually assigned on a half-point basis in the scale. The municipalities exposed to a predicted $5.75 \leq MMI < 7.5$ are initially dropped from the analysis because although the overall damages caused by the earthquake in these municipalities were small, I cannot rule out the possibility that the earthquake affected poor constructions or generated power cuts in them, affecting the benefits and costs of committing crime. Because the selection of the exact predicted intensity thresholds is to some extent arbitrary, I will examine the robustness of the results to the use of alternative intensity thresholds to define treatment and control municipalities and also to the use of the alternative method to predict earthquake intensity described in Astroza et al. (2010)¹⁰.

Figure 3.1 shows maps with (a) the predicted earthquake intensities for Chilean municipalities calculated using the method developed in Barrientos (1980) and rounded at the 0.5 points in the MMI scale and with (b) treatment and control municipalities under the default thresholds of $MMI \geq 7.5$ for treatment municipalities and $MMI < 5.75$ for control municipalities. The configuration of treatment, control and excluded municipalities under alternative earthquake intensity thresholds and calculation methods used to predict earthquake intensities are presented in figures 3.8 and 3.9 in appendix 3.C.

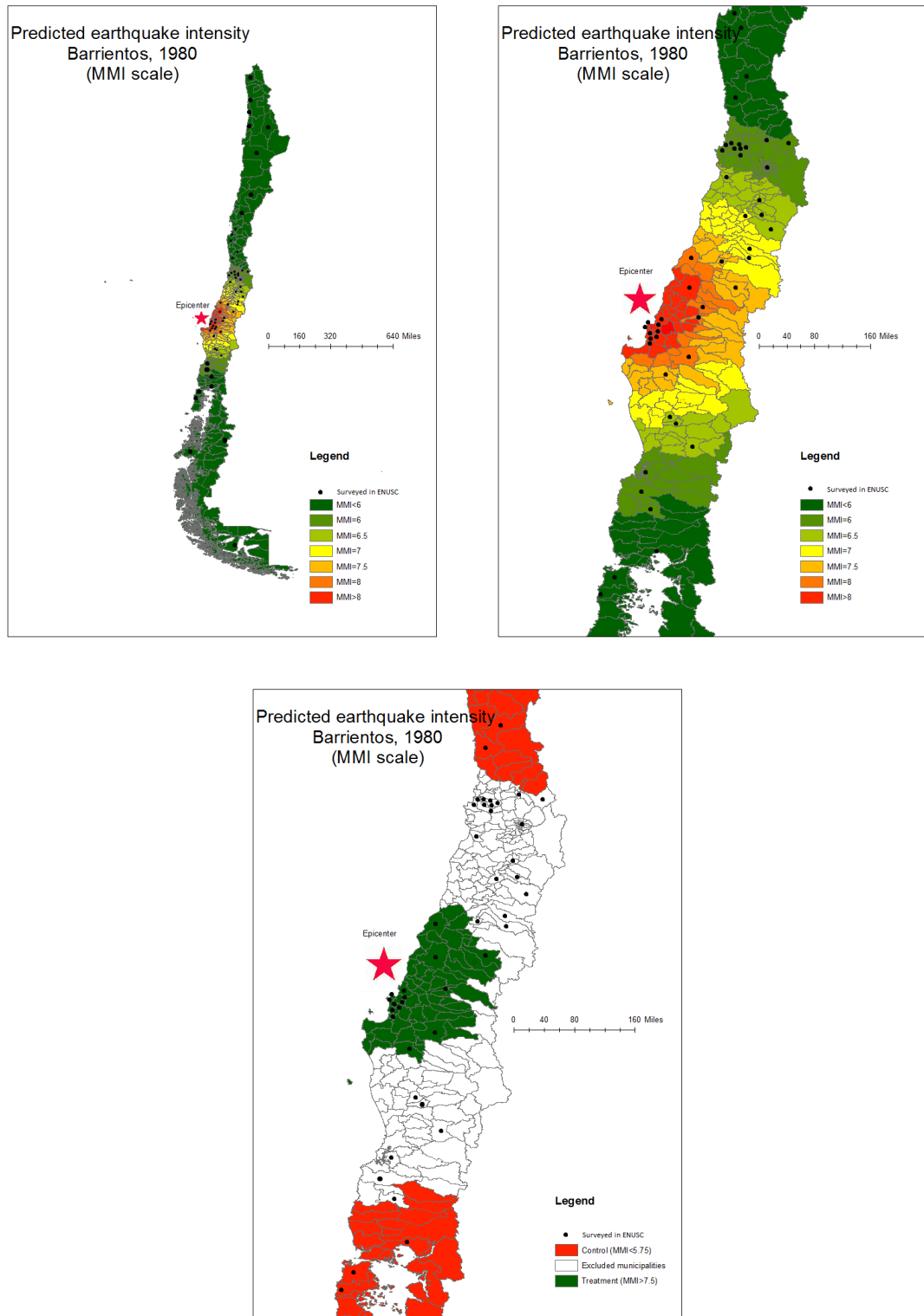
predicts the intensity of an earthquake in a given location (measured in MMI) as a function of the distance to the hypocentre and of the magnitude of the earthquake measured in M_w .

$$I_{MMI} = 1.3844M_w - 3.7355\log_{10}(DistHC) - 0.0006DistHC + 3.8461 \quad (3.1)$$

⁹The interpretation of the values in the Mercalli and MSK scales is reported in appendix 3.B.

¹⁰The paper measures MSK in 98 locations after the 2010 earthquake and estimate the MSK as a function of the distance to the closest seismic asperity. They estimate the following equation for the 2010 Chilean Earthquake:

$$I_{MSK} = 43.11 - 18.96\log_{10}(DistAs) + 0.0294DistAs \quad (3.2)$$

Figure 3.1: Predicted intensity: treatment and control areas

Note: In the maps that display the predicted earthquake intensities, the colour is assigned based on a rounding of the predicted earthquake intensity at the 0.5 points. On the other hand, the construction of the treatment and control groups of municipalities is based on whether the exact value of the predicted earthquake intensity in the municipality is above or below a certain threshold. This is the reason why for example, the municipalities exposed to a predicted earthquake intensity $7.25 \leq \text{MMI} < 7.5$ are coded as MMI 7.5 in the maps that display the predicted earthquake intensities but they are not coded as treatment municipalities in the other map.

Table 3.2: Descriptive Statistics for variables used in the analysis

	Treatment and Control before earthquake Municip. included in ENUSC data					Treatment and Control before earthquake All Treat and Contr. Municip.					All Chile (345 Munic.)				
	Treatment (16 Munic.)		Control (19 Munic.)		Diff Treat-Contr	Treatment (61 Munic.)		Control (100 Munic.)		Diff Treat-Contr	Before earth.		All periods		
	N	Mean	N	Mean		N	Mean	N	Mean		N	Mean	N	Mean	
<i>ENUSC data (household level)</i>															
Home burglary (0/1)	12,314	0.071	15,809	0.047	0.02***						74,162	0.053	177,889	0.048	
Dog (0/1)	12,314	0.399	15,812	0.410	-0.01						74,168	0.406	177,900	0.411	
Bars windows/doors (0/1)	12,314	0.470	15,812	0.442	0.03						74,168	0.540	177,900	0.549	
Safety lock (0/1)	12,314	0.311	15,812	0.235	0.08**						74,168	0.289	177,900	0.343	
Alarm (0/1)	12,314	0.073	15,812	0.060	0.01						74,168	0.097	177,900	0.113	
Share number with neigh (0/1)	12,314	0.265	15,812	0.225	0.04*						74,168	0.256	177,900	0.290	
Comm. vigilance (0/1)	12,314	0.125	15,812	0.077	0.05***						74,168	0.123	177,900	0.150	
Coord. with author. (0/1)	12,314	0.360	15,812	0.294	0.07**						74,168	0.296	177,900	0.319	
Comm. hires priv. vig. (0/1)	12,314	0.070	15,812	0.055	0.01						74,168	0.094	177,900	0.107	
<i>SPD data (municip. level)</i>															
Robbery 1,000 inhab	48	2.579	57	2.038	0.54	183	1.094	300	0.771	0.32*	1,035	1.516	2,415	1.442	
MV theft 1,000 inhab	48	0.454	57	0.918	-0.46*	183	0.179	300	0.298	-0.12*	1,035	0.513	2,415	0.645	
Larceny 1,000 inhab	48	6.021	57	6.542	-0.52	183	4.801	300	4.787	0.01	1,035	5.024	2,415	5.398	
Non-home burglary 1,000 inhab	48	2.580	57	2.309	0.27	183	2.464	300	2.204	0.26	1,035	2.452	2,415	2.583	
Home burglary 1,000 inhab	48	4.653	57	4.176	0.48	183	3.019	300	2.466	0.55*	1,035	3.466	2,415	3.499	
<i>CASEN data (municip. level)</i>															
Poverty rate	16	0.225	19	0.144	0.08***	61	0.230	89	0.126	0.10***	334	0.170	658	0.165	
Extreme poverty rate	16	0.058	19	0.035	0.02*	61	0.060	79	0.035	0.03***	324	0.046	648	0.039	
Unemployment rate	16	0.126	19	0.087	0.04***	61	0.121	79	0.085	0.04***	324	0.104	648	0.093	
Income polariz. (75% vs 25%)	16	8.313	19	7.487	0.83	61	7.175	79	7.529	-0.35	324	7.270	648	7.585	
Income polariz. (90% vs 10%)	16	21.037	19	18.708	2.33	61	17.782	79	18.703	-0.92	324	18.225	648	18.696	
Rate men between 15-29	16	0.126	19	0.122	0.00	61	0.117	89	0.111	0.01*	334	0.116	658	0.116	
Rate pop 13-25 attending educ	16	0.634	19	0.574	0.06***	61	0.583	89	0.577	0.01	334	0.576	658	0.579	
<i>Other admin data (municip. level)</i>															
Population (inhab)	48	100,398	57	118,884	-18,486	183	38,475	300	30,733	7,742	1,035	48,589	2,415	49,520	
Distance (km) to nearest city (250,000 inhab)	48	47	57	211	-164**	183	69	300	298	-230***	1,035	133	2,415	133	
% rural population	48	0.112	57	0.112	0.00	183	0.363	300	0.477	-0.11**	1,035	0.380	2,415	0.378	
Policemen per 100,000 inhab	48	214.667	57	193.754	20.91	183	178.295	300	668.927	-490.63***	1,035	318.910	2,415	328.984	
Mothers assoc. per 100 inhab.	41	0.050	51	0.024	0.03	160	0.072	277	0.061	0.01	949	0.055	2,240	0.055	
Elderly assoc. per 100 inhab.	42	0.065	51	0.050	0.02	162	0.100	277	0.092	0.01	953	0.090	2,248	0.099	
Sport clubs per 100 inhab.	42	0.119	51	0.146	-0.03	162	0.168	277	0.270	-0.10***	953	0.203	2,248	0.205	
Municipality budget per capita (2008-2013)	31	71.593	38	82.204	-10.61*	120	107.891	195	318.545	-210.65***	680	174.172	2,057	216.105	
Share aid over municipality budget (2011-2013)	48	0.017	57	0.000	0.02**	183	0.031	298	0.002	0.03***	1,031	0.014		0.014	

Note: Different data sources provide information for different periods of time. Control municipalities are those with a predicted $MMI < 5.75$ and treatment municipalities are those with a predicted $MMI \geq 7.5$, calculated following Barrientos (1980). ENUSC and CASEN surveys were not applied in all the municipalities. Descriptive statistics are provided for three different groups: (1) treatment and control municipalities included in the ENUSC survey for the years before the earthquake (2007-2009 for the SPD, ENUSC and other admin data and 2009 for the CASEN data), (2) all treatment and control municipalities for the years before the earthquake (2007-2009 for the SPD, ENUSC and other admin data and 2009 for the CASEN data) and (3) all Chilean municipalities (treatment, control and intermediate) and periods available in each data source (2007-2013 for the SPD, ENUSC and other admin data and 2009 and 2011 for the CASEN data). The values for the variable *share of reconstruction aid over municipality budget* are only reported for the years after the earthquake.

An advantage of using predicted intensity as a measure of whether a municipality is affected by the earthquake is that while this only depends on the distance to the hypocentre, the extent of human losses, economic damages or even observed earthquake intensity (which is affected by the topography of the location) at the municipality level could arguably be affected by pre-disaster factors that may influence crime costs and benefits.

Table 3.2 summarizes the data used in the study. Descriptive statistics are provided for three different samples. The first of them includes the municipalities exposed to either a predicted $MMI \geq 7.5$ or $MMI < 5.75$ that are also included in the ENUSC database, and therefore, that are used in the main analysis of the study. For this sample, the table reports the before-earthquake mean values for the variables of interest for the treatment and control groups. The second sample includes all the municipalities exposed to either a predicted $MMI \geq 7.5$ or $MMI < 5.75$ regardless of whether they are included in the ENUSC database. This is the sample of municipalities that is used in the analysis of crime data from police records and in most of the analysis of mechanisms. For this sample, the table reports the before-earthquake mean values for the treatment and control groups. The third sample includes all the Chilean municipalities regardless of their predicted MMI intensity. The table reports the mean values for the variables used in the analysis both using only the before-earthquake periods and all periods available.

The descriptive statistics for the first two samples show that before the earthquake, treatment and control municipalities were different in terms of some socioeconomic outcomes. For example, the table reveals that before the earthquake, treatment municipalities were significantly poorer, had higher rates of unemployment and lower per-capita public budgets than control municipalities. These differences between treatment and control municipalities are relevant in both the first sample (including only the municipalities surveyed in the ENUSC database, mainly urban areas) and in the second sample (that includes all treatment and control municipalities).

The pre-earthquake incidence of home burglary calculated using the ENUSC data was approximately 2.4 percentage points larger in treatment municipalities: while the probability of suffering a home burglary during the last 12 months was 4.7% in control municipalities, the 7.1% of the households living in treatment municipalities experienced a home burglary during the same period. The difference is significant at the 1%. On the other

hand, the data from police records suggest that the incidence of known to the police crime before the earthquake in treatment municipalities included in the ENUSC database was, overall, not significantly different from the incidence in control municipalities. However, some significant differences arise between treatment and control municipalities when the sample is not restricted to those municipalities included in the ENUSC database, confirming a significantly higher incidence of home burglary and robbery and a lower incidence of motor-vehicle theft in treatment municipalities before the earthquake. An interesting pattern that emerges from the comparison between crime data from the ENUSC survey and from police records is that although the exact comparison is not possible, the incidence of home burglary in police records seems much lower than in the ENUSC data. The difference could be partially explained by the fact that approximately the 50% of these offences are not reported to the police¹¹.

The information on the adoption of crime prevention measures collected in the ENUSC survey shows that overall, individuals in treatment municipalities were more likely to adopt individual and community-based crime prevention measures before the earthquake. On the other hand, the number of policemen per capita was not significantly different before the earthquake in treatment and control municipalities included in the ENUSC survey although when all treatment and control municipalities are considered, the number of policemen per capita before the earthquake in treatment municipalities was significantly lower. Finally, the strength of social life and the income inequality did not seem to differ before the earthquake in treatment and control municipalities.

3.6 Empirical Strategy

Earthquakes are natural disasters which its occurrence cannot be anticipated. However, some places are more likely to be affected by strong earthquakes. For example, areas lying in the interaction of two or more tectonic plates are more likely to suffer earthquakes of high intensity. This is indeed the case for Chile, a country with almost its entire surface lying in the border of the South-American, Nazca and Antarctic plates. Since 1900, Chile suffered 14 earthquakes of Richter magnitude equal or larger than 8 with epicentre in every Chilean region with the exception of the southern regions of Magallanes and Aysen. However, although the exact location of an earthquake cannot be considered random not

¹¹See table 3.11.

even within Chile, the timing of its occurrence can be assumed so (Cavallo et al., 2010).

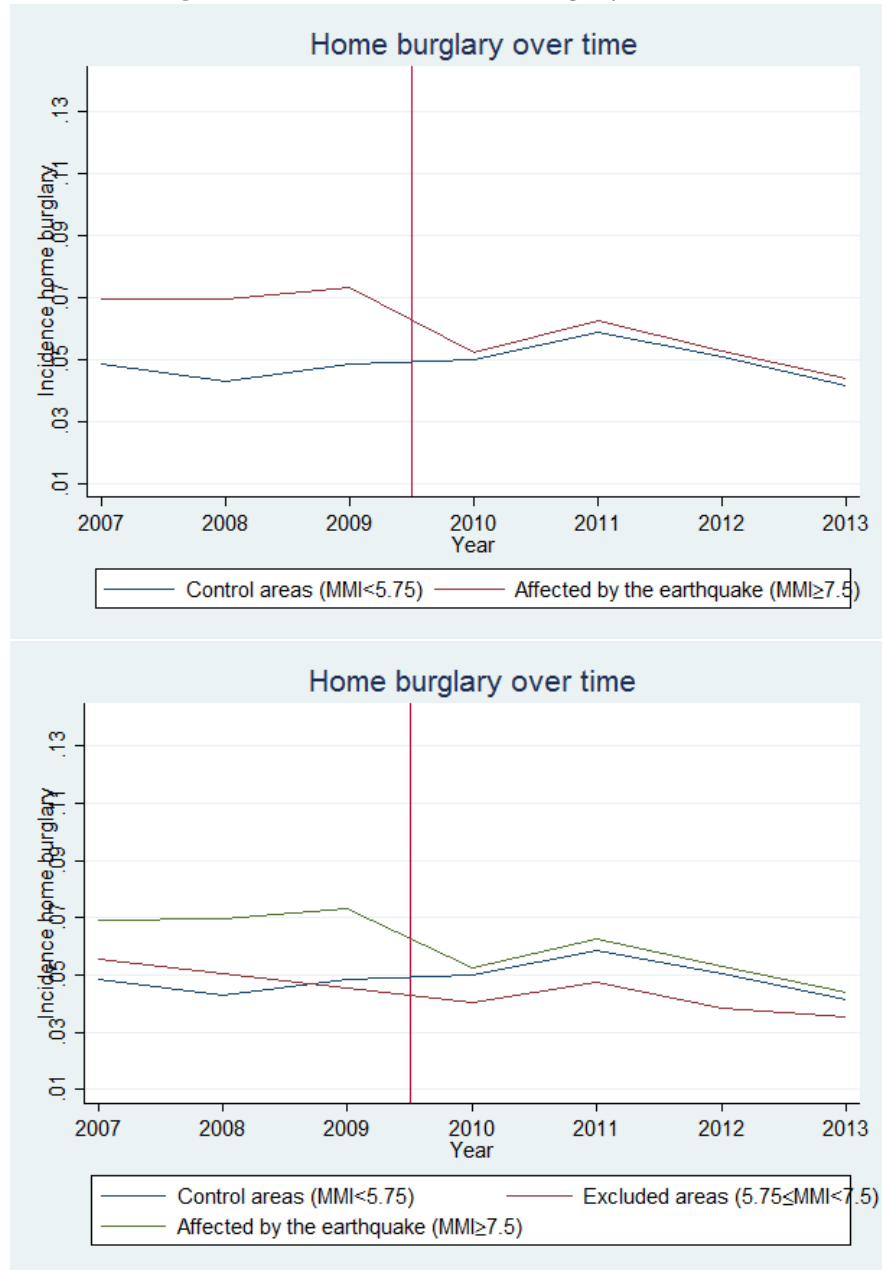
The exogenous nature of the timing in which an earthquake occurred and the impossibility to anticipate it set an ideal scenario for the use of a difference in difference strategy exploiting across-municipality and over-time variation in exposure to the earthquake for the identification of the lasting effects of exposure to the earthquake on property crime. Relying on comparing treatment and control units before and after exposure to a treatment, the difference in difference approach has been used in seminal papers to address a large variety of crucial research questions such as the effect of minimum wage on employment (Card and Krueger, 1994), the effect of school term length on student performance (Pischke, 2007) or the effect of employment protection on firms' outsourcing (Autor, 2003).

The results presented in table 3.2 suggest that treatment and control municipalities were different in terms of some socioeconomic characteristics and of the incidence of crime before the earthquake. However, the validity of the difference in difference approach in our setting does not rely on the comparability of treatment and control groups before the earthquake but on the assumption that in absence of the earthquake, crime rates in control and treatment areas would have followed the same trajectory over time. This identifying condition can be partially tested through assessing whether before the earthquake, property crime rates in areas close and far away from the epicentre followed the same trend over time. If the evolution over time of crime rates was similar in treatment and control municipalities before the earthquake, it would be reasonable to assume that if the earthquake had not occurred, areas next to and far from the hypocentre would have followed the same crime trend over time during all the period studied. Figure 3.2 plots the evolution over time for the period 2007-2013 of the incidence of home burglary by level of exposure to the earthquake. A visual inspection of the latter figure suggests that although the levels are different, the evolution of the incidence of home burglary over time before the earthquake was the same in the areas next to the hypocentre that are defined as treatment municipalities and in the areas far from the hypocentre that are defined as control municipalities. On the other hand, figure 3.2 also suggests that before the 2010 earthquake, the incidence of home burglary in those intermediate municipalities excluded from the analysis was following a different trend over time. The lack of pre-earthquake parallel trends in these municipalities is also confirmed empirically¹², implying that the

¹²To test this hypothesis, I estimate a leads and lags model and test the joint significance of the lead variables. The F-test is significant at the 10% confidence level.

effect of the earthquake on crime rates in these intermediate municipalities cannot be reliably estimated using a difference in difference strategy.

Figure 3.2: Incidence of home burglary over time



For the identification of the effect of the earthquake on property crime dynamics, I estimate two models using the ENUSC database formed of seven repeated cross sections of households and omitting from the sample the households living in intermediate municipalities. Following Autor (2003), I first estimate a leads and lags model:

$$Burglary_{imt} = \alpha_m + \sum_{\tau=-q}^{-1} \beta_{\tau}(Year_t \times Earthquake_m)_{m\tau} + \sum_{\tau=0}^r \beta_{\tau}(Year_t \times Earthquake_m)_{m\tau} + Year_t + \mu_{imt} \quad (3.3)$$

where $Burglary_{imt}$ is a dummy variable equal to 1 if household i in municipality m and in year t has suffered a home burglary in the last 12 months and 0 otherwise. $Year$ is a vector of year dummy variables, α_m is a vector of municipality dummies, and $(Year \times Earthquake)_{mt}$ is a vector of variables constructed as the interaction of the dummy variable $Earthquake_m$ that is equal to 1 if the municipality m was exposed to a predicted $MMI \geq 7.5$ and 0 otherwise, with each year dummy. These interaction variables are known in the literature as the lead and lag variables. In our specification, the lead variables are the interaction between year and earthquake exposure for the years before the earthquake (from period $\tau = -q$ to period $\tau = -1$). The lag variables are the interaction between year and earthquake exposure for the years after the earthquake (from period $\tau = 0$ to period $\tau = r$). The coefficients of the lead and lag variables yield the differential variation in the home burglary rate in treatment and control municipalities in the year of interest relative to 2009, the last year before the earthquake and the omitted category in the regression specification. The coefficients of the lead and lag variables estimated in equation 3.3 pursue a double objective. First, the estimated coefficients for the lead variables provide an empirical test for the parallel trends condition. If these coefficients are small and statistically indistinguishable from 0, the home burglary rate in treatment and control municipalities was arguably following the same trend before the earthquake. Second, if the coefficients for the lead variables are statistically indistinguishable from 0, the coefficients for the lag variables yield the effect of the earthquake on the incidence of burglary over time, providing information on the dynamics and persistence of this effect.

Second, I also estimate the following regression:

$$Burglary_{imt} = \alpha_m + \beta(Earthquake \times POST)_{mt} + Year_t + u_{imt} \quad (3.4)$$

where $(Earthquake \times POST)_{mt}$ is an interaction term of the variable $Earthquake$ that indicates whether municipality m was exposed to a predicted $MMI \geq 7.5$ and the variable $POST$, that is equal to 1 for those periods after the earthquake. The parameter β yields

the pooled effect of exposure to the earthquake on the incidence of home burglary over the period of interest (2010-2013) relative to municipalities not directly affected by the earthquake. Following the recommendation of Angrist and Pischke (2008) for difference in difference estimations with several pre- and post-treatment periods, I clustered the standard errors at the municipality level.

Although the earthquake plausibly caused negligible *direct* economic damage in control municipalities, it is not possible to rule out the possibility that the earthquake affected *indirectly* economic outcomes in these municipalities. For example, the central government might have allocated some investments planned for municipalities not directly affected by the earthquake to the reconstruction of the most devastated municipalities. I discuss in section 3.7 the existence of *indirect* effects of the earthquake in control municipalities and the extent to which these *indirect* effects could affect the estimates reported in this study.

3.7 Results

Table 3.3 presents the main results of the study. Column 1 reports the estimates for equations 3.3 and 3.4 when the treatment group is defined as those households living in municipalities exposed to a predicted $MMI \geq 7.5$; and the control group is defined as those households living in municipalities exposed to a predicted $MMI < 5.75$. As described in the previous section, those households living in municipalities exposed to an intensity $5.75 \leq MMI < 7.5$ are excluded from the regression. The results of the leads and lags analysis reported in column 1 are also displayed graphically in figure 3.3. Columns 2-6 report the estimates for equations 3.3 and 3.4 when alternative earthquake intensity thresholds and the alternative method developed by Astroza et al. (2010) to predict earthquake intensities are used to define treatment and control municipalities. The results of these analyses are also displayed graphically in figure 3.11 in appendix 3.E.

One of the advantages of the leads and lags approach is that it provides a direct test for the parallel trends condition in difference in difference models with more than one pre-treatment period. This condition would be satisfied if the coefficients that measure the year-specific effects of the earthquake on crime the years before the earthquake (the leads) are small and not statistically significant.

The estimates for the lag variables reported in table 3.3 show that for every threshold

used to define treatment and control groups, the coefficients for the effect of the earthquake for the years before its occurrence are very small and largely insignificant. On the other hand, the table shows that the coefficients that measure the effect of the earthquake on home burglary (the lag variables) are negative, large and statistically significant at conventional confidence levels in the majority of the specifications. Overall, the results suggest that the earthquake decreased significantly the incidence of home burglary the year of the earthquake in areas close to the hypocentre relative to those areas far away from it. The magnitude of this effect on the probability of experiencing a home burglary during the last 12 months ranges between 1.1 and 2.1 percentage points (equivalent to a 16%-30% reduction in the prevalence of home burglary relative to the last pre-earthquake period in treatment municipalities), depending on the intensity threshold used to define control and treatment municipalities. Furthermore, the effect of the earthquake remained stable during the 4 post-earthquake years studied, confirming the persistence of this effect over this period. Although the exact magnitude and level of significance for the year-effect estimates vary with the definition of the treatment and control groups, the coefficients are consistently negative and the pooled effect over the period of interest is statistically significant at the 5% in all the specifications, highlighting that the results are robust to the use of different predicted intensity thresholds to define treatment and control municipalities and to the use of the method developed by [Astroza et al. \(2010\)](#) to predict earthquake intensities.

A more detailed look at how the magnitude of the effect varies when different thresholds are used suggests that the smaller (larger) the distance to earthquake hypocentre (predicted intensity) threshold used to define the treatment group, the larger and more significant the effect of the earthquake is. In this sense, for example, the estimates in column 1 are larger in absolute value than those reported in column 2. Similarly, the larger (smaller) the distance (predicted intensity) threshold used to define the control group, the larger and more significant the effect of the earthquake is. The latter is illustrated by the fact that estimates in columns 1 and 2 are larger than those in columns 4 and 5. These results suggest that the higher the earthquake intensity exposed to, the larger the effect on home burglary. Furthermore, they also cast doubts on whether municipalities in the limit between control and intermediate areas could have been somehow affected by the earthquake and suggest that the use of longer distances from the hypocentre to define control

municipalities could be more convenient.

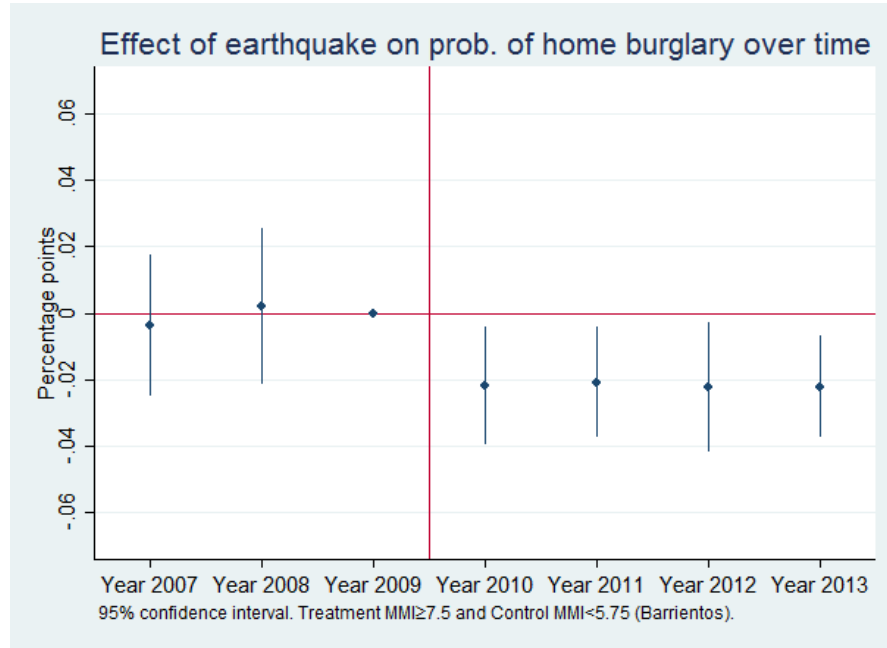
The results reported in columns 1-8 of table 3.8 in appendix 3.A confirm that the main findings are robust to the inclusion of municipality time trends in the regressions. Furthermore, the estimates provided in columns 4 and 8 show that the effect of the earthquake on home burglary is also robust to the exclusion of households living in municipalities that were affected by the tsunami, suggesting that the impact of the earthquake is not confounded by the effects of the tsunami in some earthquake affected municipalities. Finally, as expected, the inclusion of households living in intermediate municipalities as a separate group in the regression hardly changes the magnitude of the estimates for the treatment group. Indeed, although the lack of parallel trends requires to take with caution the estimates for the intermediate municipalities, the smaller but negative and statistically significant coefficient for this group relative to control municipalities suggests the possibility that the earthquake may have also decreased the prevalence of property crime in intermediate municipalities.

Finally, although the direct effect of the earthquake in control municipalities was likely negligible, I cannot rule out the existence of indirect effects of the earthquake, ultimately affecting crime in these municipalities. For example, the central government might have allocated some investments planned for municipalities not directly affected by the earthquake to the reconstruction of the most devastated municipalities, potentially affecting crime in control municipalities. If the effect of the earthquake in control municipalities had the same direction (although smaller in magnitude) than in treatment municipalities, the effect of the earthquake on property crime estimated in this section should be interpreted as a lower bound for the true effect. On the other hand, if the effect of the earthquake in control municipalities had the opposite direction than the effect in treatment municipalities, the coefficients estimated in this study would overestimate the true effect. Although I cannot reject any of the last two hypotheses, figure 3.2 shows a sharp break in the crime trends in treatment municipalities the year of the earthquake and a smooth trend in control municipalities the same year, suggesting that if any, the indirect effect of the earthquake on crime in control municipalities would be small.

Table 3.3: Effects of the earthquake on home burglary (ENUSC data): Leads and lags analysis and pooled effects for the period 2007-2013

	(1) Home burglary (0/1)	(2) Home burglary (0/1)	(3) Home burglary (0/1)	(4) Home burglary (0/1)	(5) Home burglary (0/1)	(6) Home burglary (0/1)
Specif. A: Leads and Lags						
<i>Lead var. (Parallel trends)</i>						
Earthquake × Year 2007	-0.004 (0.010)	0.001 (0.010)	-0.001 (0.010)	0.000 (0.010)	0.004 (0.010)	-0.011 (0.009)
Earthquake × Year 2008	0.002 (0.011)	0.004 (0.011)	0.002 (0.009)	0.001 (0.011)	0.004 (0.010)	-0.003 (0.009)
<i>Lag var. (Year-based effects)</i>						
Earthquake × Year 2010	-0.022** (0.009)	-0.016 (0.010)	-0.018** (0.009)	-0.016** (0.008)	-0.011 (0.009)	-0.016** (0.007)
Earthquake × Year 2011	-0.021** (0.008)	-0.016* (0.008)	-0.018* (0.009)	-0.013 (0.008)	-0.008 (0.008)	-0.016** (0.008)
Earthquake × Year 2012	-0.022** (0.010)	-0.017* (0.010)	-0.017 (0.010)	-0.018** (0.009)	-0.013 (0.009)	-0.012 (0.009)
Earthquake × Year 2013	-0.022*** (0.008)	-0.016* (0.008)	-0.019** (0.008)	-0.017** (0.007)	-0.011 (0.008)	-0.019** (0.008)
Specif. B: Pooled effect						
Earthquake × Post	-0.021*** (0.005)	-0.018*** (0.005)	-0.018*** (0.006)	-0.017*** (0.005)	-0.013** (0.005)	-0.018*** (0.006)
Observations	67,540	70,814	68,878	81,276	84,550	159,259
Sh. burglary (treatment areas)	0.071	0.067	0.065	0.071	0.067	0.065
Treatment areas						
MMI/MSK	≥ 7.5	≥ 7	≥ 7	≥ 7.5	≥ 7	≥ 7
Km hypocentre/asperity	≤ 180	≤ 239	≤ 124	≤ 180	≤ 239	≤ 124
Control areas						
MMI/MSK	< 5.75	< 5.75	< 4.9	< 6	< 6	< 5.75
Km to hypocentre/asperity	> 473	> 473	> 250	> 415	> 415	> 170
Intensity prediction method						
	Barrientos MMI/hypocentre	Barrientos MMI/hypocentre	Astroza MSK/asperity	Barrientos MMI/hypocentre	Barrientos MMI/hypocentre	Astroza MSK/asperity

Note: The table reports the estimates at the household level for the effect of the earthquake on home burglary over time using the ENUSC database and different predicted intensity thresholds to define treatment and control municipalities and methods to predict earthquake intensity. Specification A corresponds to the leads and lags model (equation 3.3). It yields the year-based effect of the earthquake during the period of interest. Specification B corresponds to the pooled effect difference in difference model (equation 3.4). It measures the average effect of the earthquake over the post-earthquake period of interest. Lead and lag variables are not included in specification B and the effect of interest is captured by an interaction between the dummy variables that capture whether the household lives in a municipality affected by the earthquake and whether the household is interviewed after the earthquake. The mean of the dependent variable is provided for the treatment areas before the earthquake. Standard errors clustered at the municipality level. ***p<0.01; **p<0.05; *p<0.1.

Figure 3.3: Effect of the earthquake on home burglary over time

3.8 Additional Analysis: Known to the Police Crime Data

This section uses the SPD database that includes yearly and monthly crime and apprehension data from police records to conduct the following analyses. First, I check the robustness of the results presented in section 3.7 to the use of a different data source and a longer pre-earthquake period, expanding also the analysis to other types of property crime. Second, I examine whether the social chaos and the episodes of looting that occurred in the aftermath of the earthquake were accompanied by sharp increases in the incidence of property crimes reported to the police. Third, I test whether within 30 days from the earthquake and in a context of looting, army deployment and the enactment of a curfew, the number of individuals apprehended by the police raised in treatment municipalities.

3.8.1 The Effect of the Earthquake on Known to the Police Property Crime

Using the SPD database, I estimate equations 3.5 and 3.6 using the yearly incidence of home burglary, non-home burglary, larcenies, motor-vehicle theft and robbery per 1,000 inhabitants as dependent variables:

$$Crime_{mt} = \alpha_m + \sum_{\tau=-q}^{-1} \beta_{\tau}(Year \times Earthquake)_{mt} + \sum_{\tau=0}^r \beta_{\tau}(Year \times Earthquake)_{m\tau} + Year_t + \mu_{mt} \quad (3.5)$$

$$Crime_{mt} = \alpha_m + \beta(Year \times Earthquake)_{mt} + \gamma X_m + Year_t + u_{mt} \quad (3.6)$$

where $Crime_{mt}$ is the incidence per 1,000 inhabitants of each specific type of property crime in the municipality m in year t . The models are estimated using OLS and standard errors are clustered at the municipality level. Note that equations 3.5 and 3.6 are similar to equations 3.4 and 3.3 although this section uses crime data available at the municipality level and therefore, the regressions are estimated using municipalities as the unit of analysis.

The estimation of equations 3.5 and 3.6 for every type of crime is conducted using two different samples. The first of them includes the period 2007-2013, which covers the years included in the ENUSC data used in section 3.7. The analysis conducted with this first sample examines the robustness of the main results to the use of a different source of data, utilizing the same time period employed in the main analysis. The second sample includes crime data for a wider pre-earthquake period, covering the years 2003-2013. The analysis of the latter sample yields information on whether the parallel trends condition still holds when a longer pre-earthquake period of time¹³ is incorporated.

The results of the analyses using these two samples are reported in table 3.4. Overall, the estimates are consistent with those obtained in the main analysis conducted in section 3.7. The coefficients reported in columns 1 and 2 suggest that earthquake decreased significantly home burglary the year of the earthquake. The effect remained significant 4 years after the earthquake although the magnitude of the effect was smaller. The estimates displayed in columns 3 and 4 show that the earthquake reduced the incidence of larceny the year of the earthquake and the magnitude of this effect remained relatively stable and statistically significant over all the period studied. The coefficients for the lag variables that measure the effects of the earthquake on non-home burglary are reported in columns 5 and 6. They are consistently negative although only statistically significant for the first post-earthquake year. Note however that the pooled effect of the earthquake on non-

¹³The first year for which the SPD database includes data for all the types of crime analysed in 2003.

home burglary over the period of interest is negative and statistically significant at the 1%. The results reported in columns 7 and 8 reveal that unlike for the previous types of crime, the earthquake did not seem to affect the incidence of motor-vehicle theft. The results for the incidence of robbery are more ambiguous. While none of the coefficients for the lag variables reported in columns 9 and 10 is statistically significant at the 10%, the pooled effect of the earthquake on robbery over the period of interest is negative and statistically significant in the sample that only includes the period 2007-13 and negative but statistically indistinguishable from 0 at conventional confidence levels when the full period 2003-2013 is analysed.

The results of the F-test for the lead variables suggest that although figures 3.4 and 3.5 reveal some differences in the pre-earthquake evolution of the known to the police incidence of home burglary, non-home burglary, larceny and motor-vehicle theft in treatment and control municipalities, these differences are not statistically significant in any of the two time periods used in the analysis. On the other hand, the results of this test show that the pre-earthquake evolution over time of the incidence of robbery is significantly different in treatment and control municipalities, casting doubts on the estimates provided in columns 9 and 10 of table 3.4.

One potential concern when interpreting the coefficients reported in table 3.4 is that unlike the ENUSC database, the SPD database only includes those offences reported to or unmasked by the police. Therefore, the SPD database misses those crimes that were neither reported to nor unmasked by the police. The reporting error in police records may generate two problems. First, a substantial share of crimes unknown to the police would lead to large standard errors. Second, if the share of crimes that is unknown to the police is affected by the earthquake, the models would yield biased estimates of the effect of the earthquake on *true* property crime rates. For example, at the limit, the results discussed in table 3.4 might be explained by an effect of the earthquake on the share of crimes that is reported to or unmasked by the police rather than by an effect of the earthquake on *true* crime rates.

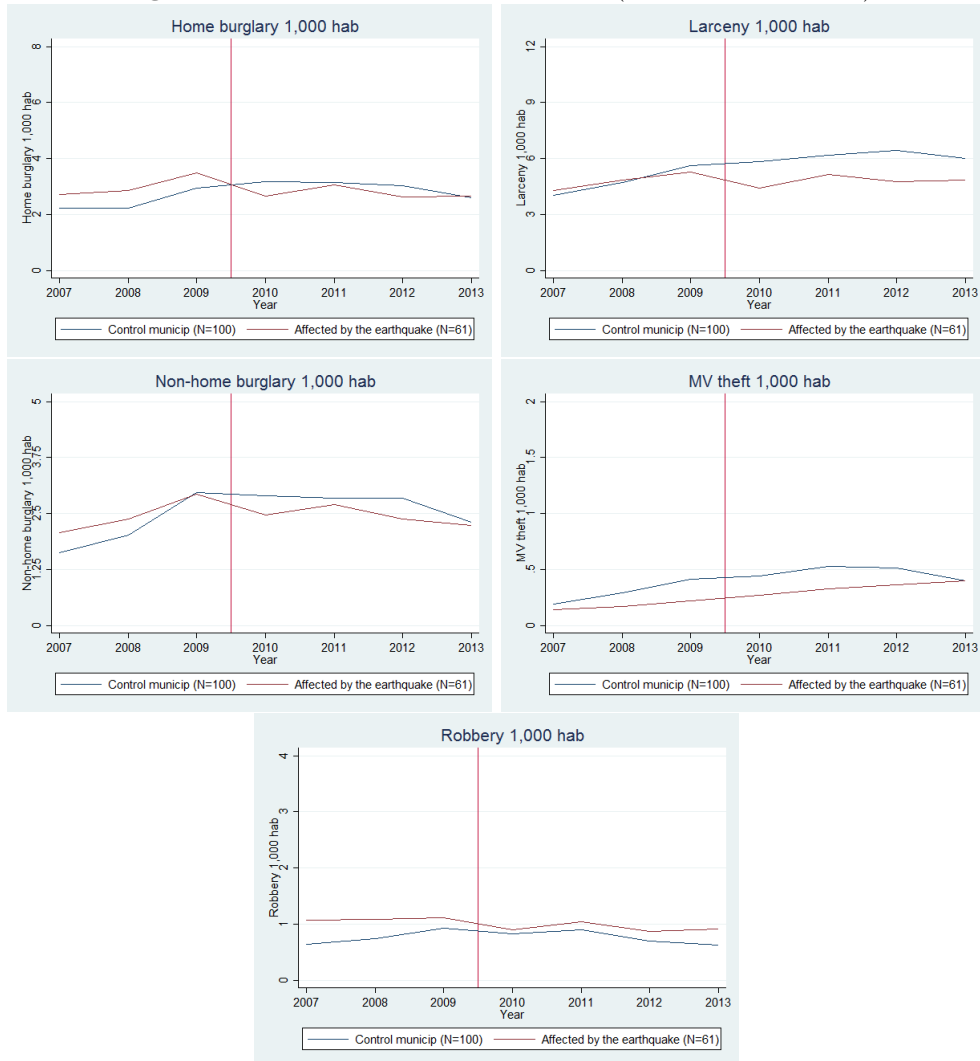
Table 3.4: Effects of earthquake exposure on property crime (2007-2013): SPD data

	(1) Home burglary 1,000 inhab	(2) Home burglary 1,000 inhab	(3) Larceny 1,000 inhab	(4) Larceny 1,000 inhab	(5) Non-home burgl. 1,000 inhab	(6) Non-home burgl. 1,000 inhab	(7) MV theft 1,000 inhab	(8) MV theft 1,000 inhab	(9) Robbery 1,000 inhab	(10) Robbery 1,000 inhab
Specif. A: Leads and lags										
<i>Lag var. (Year-based effects)</i>										
Earthquake \times Year 2010	-1.080*** (0.222)	-1.080*** (0.229)	-1.113*** (0.255)	-1.113*** (0.262)	-0.387* (0.225)	-0.387* (0.231)	0.020 (0.045)	0.020 (0.046)	-0.115 (0.091)	-0.115 (0.093)
Earthquake \times Year 2011	-0.625** (0.259)	-0.625** (0.267)	-0.698** (0.320)	-0.698** (0.329)	-0.099 (0.276)	-0.099 (0.284)	-0.003 (0.081)	-0.003 (0.084)	-0.033 (0.089)	-0.033 (0.089)
Earthquake \times Year 2012	-0.955*** (0.295)	-0.955*** (0.304)	-1.359*** (0.439)	-1.359*** (0.452)	-0.423 (0.299)	-0.423 (0.308)	0.041 (0.079)	0.041 (0.081)	-0.014 (0.081)	-0.014 (0.084)
Earthquake \times Year 2013	-0.478* (0.283)	-0.478* (0.291)	-0.828** (0.359)	-0.828** (0.370)	-0.031 (0.274)	-0.031 (0.282)	0.188*** (0.060)	0.188*** (0.062)	0.093 (0.077)	0.093 (0.079)
Specif. B: Pooled effect										
Earthquake \times Post	-0.508*** (0.174)	-0.786*** (0.185)	-1.669*** (0.288)	-1.327*** (0.293)	-0.491*** (0.157)	-0.529*** (0.166)	-0.017 (0.074)	-0.011 (0.060)	-0.096 (0.067)	-0.158** (0.077)
Mean dep. var in Treatment municip 2009	3.640	3.640	5.161	5.161	3.052	3.052	0.246	0.246	1.105	1.105
Observations	1,769	1,127	1,767	1,127	1,769	1,127	1,769	1,127	1,767	1,127
Treatment municip.	61	61	61	61	61	61	61	61	61	61
Control municip.	100	100	100	100	100	100	100	100	100	100
Sample (Years)	2003-13	2007-13	2003-13	2007-13	2003-13	2007-13	2003-13	2007-13	2003-13	2007-13
Pre-earthq. trends										
F-test: lead variables										
$H_0 : \beta_{T=-q} = \dots = \beta_{T=-1} = 0$	1.707	0.340	1.521	1.938	1.741	1.606	1.571	1.941	2.721**	3.641**

Note: The table reports the estimates at the municipality level for the effect of the earthquake on different types of property crime using data from crime records. Specification A corresponds to the leads and lags model (equation 3.5). It yields the year-based effect of the earthquake during the period of interest. Specification B corresponds to the pooled effect difference in difference model (equation 3.6). It measures the average effect of the earthquake over the post-earthquake period of interest. Lead and lag variables are not included in specification B and the effect of interest is captured by an interaction between the dummy variables that capture whether the municipality is affected by the earthquake and whether the observation corresponds to a year after the earthquake. For each type of crime and specification, two samples are used. The first includes only the period 2007-2013, which is the period used in the analysis of the ENUSC data. The second sample includes the period 2003-2013, using all the pre-earthquake years for which the SPD data is available. A test for the common trends assumption is reported for every estimation. For this, I use an F-test to examine the joint significance of the lead variables. The mean of the dependent variable is provided for the treatment areas in the last year before the earthquake (2009). Standard errors clustered at the municipality level. ***p<0.01; **p<0.05; *p<0.1.

I explore this hypothesis using information available in the ENUSC survey on whether households report crimes to the police and estimating a difference in difference model in which the dependent variable is the share of larcenies, motor-vehicle theft, robbery and home burglary that is reported to the police. The results of this analysis, conducted at the regional level, are reported in table 3.11 in appendix 3.D¹⁴. Both the regression analysis and the visual inspection of figure 3.10 suggest that the earthquake does not systematically affect the share of crime that is reported to the police. Nonetheless, and even if the earthquake does not affect the probability of reporting crime to the police, the fact that approximately 50% of the home burglaries and robberies and 75% of the larcenies are not reported to the police introduces measurement error in the dependent variable, leading to wider standard errors for the coefficients reported in table 3.4.

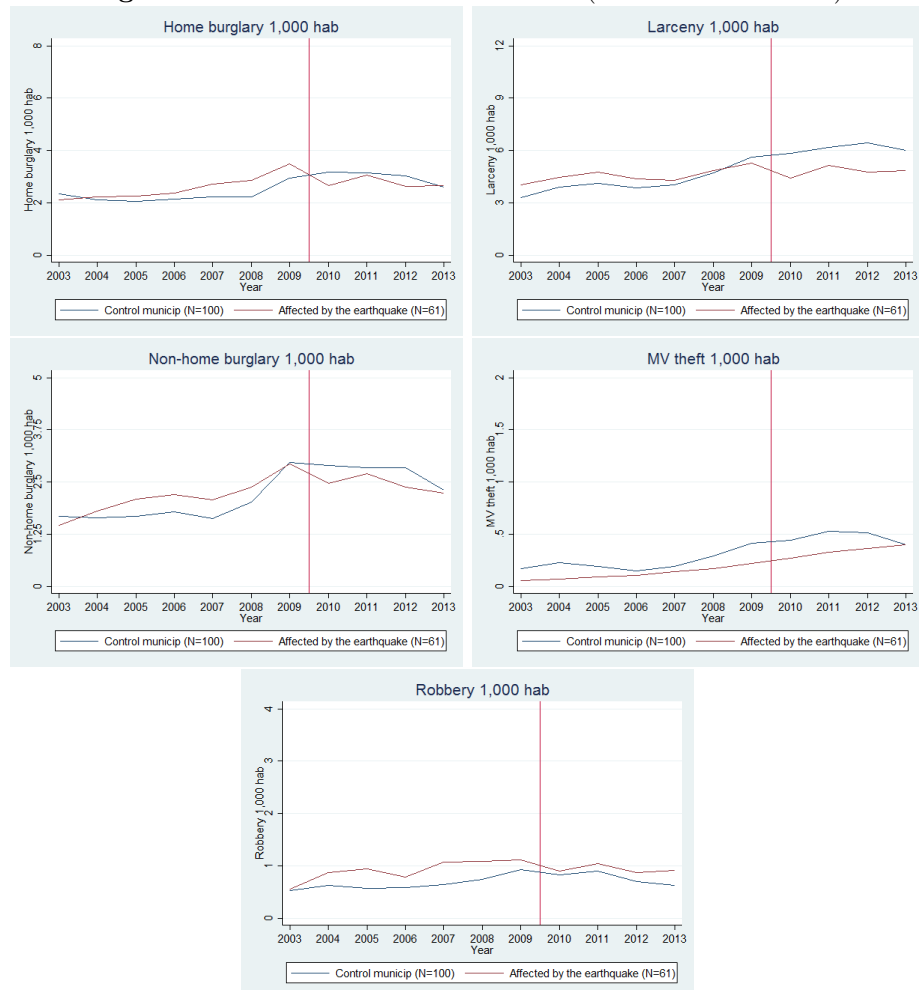
Figure 3.4: Incidence of crime over time (SPD data 2007-2013)



¹⁴The analysis is conducted at the regional level because the ENUSC survey does not provide the location at the municipality level for most of the offences (larceny, motor-vehicle theft and robbery) that are used to construct the dependent variable.

Overall, the results presented in this section are consistent with the findings of the analysis conducted in section 3.7 and show that the earthquake led to a lasting reduction in the incidence of home burglary. Furthermore, the results on the different types of property crimes also exclude the possibility that rather than decreasing property crime, the earthquake simply displaced criminals from engaging in burglary to commit other types of property crime. The latter hypothesis, studied by [Bell et al. \(2014\)](#) in the context of the 2011 London riots, could be relevant if judges increased the severity of sentencing for criminals committing burglaries in areas affected by the earthquake or if criminals falsely perceived more severe sentencing for these crimes. Although the criminal law did not change following the earthquake, the social awareness and media coverage of the looting events may have induced judges in these areas, at least temporally, to increase the severity of sentencing for burglary. However, the fact that the reduction in crime rates seems to operate over different types of property crime precludes the hypothesis that the earthquake simply displaced crime from burglary to other types of property crime.

Figure 3.5: Incidence of crime over time (SPD data 2003-2013)



3.8.2 Crime and Punishment in the Aftermath of the Earthquake

Some treatment municipalities experienced looting episodes, the enactment of a curfew and the deployment of the army during the two weeks that followed the earthquake. Through incapacitating criminals or providing a first contact with crime for some looters, the deployment of the army, the curfew and the looting episodes could have affected the incidence of crime in treatment municipalities in a lasting way.

This subsection investigates crime and apprehension in the aftermath of the earthquake through estimating the effect of exposure to the earthquake on the incidence of different crimes and on the apprehension rate the month of the earthquake and one month after the earthquake using the monthly SPD data. The dependent variables in these regressions are the change in the incidence of crime between either the month of the earthquake (February 2010) or the first month after the earthquake (March 2010) and the last month before the earthquake (January 2010). The regression includes as control variables the population of the municipality and the incidence of crime in the last month before the earthquake.

The results of these analyses are reported in table 3.5. They suggest that property crime did not increase sharply in the aftermath of the earthquake. Rather, the known to the police incidence of home burglary, robbery and larceny one month after the earthquake was significantly lower in earthquake affected municipalities. Although these results could be surprising, [Grandón et al. \(2014\)](#) suggest that the most prevalent type of property crime after the earthquake was group looting towards large supermarkets and shops that although involved many people, in terms of numbers of offences reported to the police might be small. Furthermore, the same study highlights that the looting of houses or small shops was an extremely rare event in the aftermath of the earthquake. In any case, the latter estimates should be interpreted with caution because crime data from police records aggregated at the monthly level might not be the most suitable for this analysis. First, the cost of reporting to the police an offence might be larger the days followed the earthquake due to institutional collapse, potentially leading to an underestimation of the short-term effect of the earthquake on *true* crime rates. Second, the effect of the earthquake on crime might be restricted to a few days or hours after its occurrence and before the deployment of the army. However, the aggregation of the crime data at the monthly level may not be adequate to assess the very short-term effects of the earthquake.

Table 3.5: Impact estimates (OLS): Short-term effects of the earthquake on different types of property crimes and on individuals apprehended (SPD data)

	Δ Home burglary (per 1,000 inhab)	Δ Larceny (per 1,000 inhab)	Δ Non-home burglary (per 1,000 inhab)	Δ Motor-vehicle thefts (per 1,000 inhab)	Δ Robbery (per 1,000 inhab)	Δ Apprehended (per 1,000 inhab)
<i>Sample A: March 2010 - Jan 2010</i>						
Earthquake municip.	-0.098*** (0.028)	-0.211*** (0.052)	0.039 (0.051)	-0.004 (0.006)	-0.026** (0.012)	-0.155** (0.066)
<i>Sample B: Feb 2010 - Jan 2010</i>						
Earthquake municip.	-0.083 (0.054)	-0.126* (0.065)	-0.101 (0.088)	-0.009 (0.009)	0.012 (0.017)	0.067 (0.074)
Observations	161	161	161	161	161	161
Av. rate Jan 2010 (Treat mun)	0.328	0.527	0.215	0.057	0.138	0.542

Note: The regressions estimated use monthly data from police records (SPD database) and OLS methods to estimate at the municipality level the short term effects of the earthquake on property crime and on individuals apprehended. The equation estimated is $\Delta Y = \beta_0 + \beta_1 \text{Earthquake} + \beta_2 Y + \mu$ where the dependent variable ΔY is the difference in crime rates/individuals apprehended between March 2010 (the month after the earthquake) and January 2010 (the last month before the earthquake) in sample A and the difference in crime rates/individuals apprehended between February 2010 (the month of the earthquake) and January 2010 in sample B. Y measures the crime rate/number of people apprehended in January 2010. Municipalities exposed to a predicted $5.75 \leq MMI < 7.5$ are excluded from the analysis. Robust standard errors in parentheses. ***p<0.01; **p<0.05; *p<0.1.

The results reported in the last column in table 3.5 highlight that far from increasing, the number of individuals apprehended per 1,000 individuals decreased in the aftermath of the earthquake relative to control municipalities. These results dismiss the possibility that the contraction in property crime rates in earthquake affected municipalities is driven by a higher rate of incarceration in these municipalities as a consequence of the curfew and the deployment of the army.

One explanation for the reduction in incarceration rates and in the known to the police incidence of some types of property crimes in the aftermath of the earthquake could be the deterring effect of the army deployment and of the curfew. If so, through temporarily increasing the cost of crime, the presence of the army and the curfew may have persistent effects on the incidence of crime. This hypothesis is explored in section 3.9 as a potential mechanism for the lasting reduction in property crime after the earthquake.

3.9 Analysis of Mechanisms

Natural disasters are complex phenomena that may influence the benefits and costs of crime through many channels. This section discusses the relevance of some of the most evident ones. However, it is beyond the scope of the study to comprehensively examine every individual path through which the earthquake could have reduced the incidence of property crime over the post-earthquake period studied.

The lasting reduction in property crime after the earthquake is consistent with the predictions of the informal guardianship theory. The latter argues that natural disasters are generally followed by altruistic behaviours that strengthen community links and co-operation, increasing the provision of informal guardianship in damaged communities and therefore the costs of crime. The theory concludes that the rises in the levels of informal guardianship offset the potential perverse effects of disasters on crime caused by their negative impact on other crime determinants such as the capacity of the police to enforce the law.

In order to test the informal guardianship channel, I estimate equation 3.4 at the household level using as dependent variables the information collected in the ENUSC survey on adoption of different household and community-based measures to prevent crime. The results of this analysis are displayed in table 3.6 and show that the earthquake boos-

ted the provision of informal guardianship by households, mainly through the adoption of community-based measures such as creating community alarms or sharing telephone numbers with neighbours. Furthermore, the estimates reported in column 10 of table 3.8 in appendix 3.A suggest that the drop in the incidence of home burglary was more than twice among households living in municipalities affected by the earthquake that increased the provision of community-based strategies to prevent crime than among households living in earthquake affected municipalities that did not increase it, although the effects are not statistically different from each other at conventional significance levels¹⁵. Although the rise in the incidence of community-based measures to prevent crime among treatment municipalities was probably not random and therefore the results of this analysis should not be interpreted as causal, the estimates point to this mechanism as an important path through which the earthquake may have decreased crime. Also in line with the informal guardianship theory, the coefficients reported in columns 9-11 of table 3.6 suggest that, overall, the earthquake increased the number of community-based organizations. This finding is consistent with a positive effect of the earthquake on the strength of community life. Finally, qualitative studies analysing social dynamics in the aftermath of the earthquake remark the widespread prevalence of pro-social, altruistic and organized behaviour in communities affected by the earthquake during the days that followed the natural disaster (Grandón et al., 2014; Larrañaga and Herrera, 2010b).

However, the rise in the adoption of community-based crime prevention measures and the provision of informal guardianship could be also driven by an increase in the perceived risk of crime in communities affected by the earthquake. If so, the rise in the provision of informal guardianship and the drop in crime could have happened even in the absence of any effect of the earthquake on social capital. Although 10 months after the earthquake the perception of crime was not significantly different in earthquake affected and unaffected areas¹⁶, it is likely that the extensive media coverage of the looting events and the power

¹⁵Column 10 of table 3.8 reports the estimation using two separate treatment groups. The first treatment arm includes those municipalities affected by the earthquake that increased the provision of community-based strategies to prevent crime after the earthquake. A municipality is considered to have increased the provision of community-based strategies to prevent crime when the average number of community-based measures to prevent crime (including sharing telephone numbers with neighbours, organizing community vigilance, coordinating with local authorities for the provision of security and hiring private vigilance) adopted in post-earthquake years in the municipality is higher than in pre-earthquake years. The second treatment group includes those municipalities affected by the earthquake that did not increase the provision of community-based strategies to prevent crime after the earthquake.

¹⁶This analysis, conducted using the ENUSC data, is not reported in the paper but it is available upon request.

cuts increased the perception of crime in earthquake affected communities even when the looting of houses and small business was very rare in the aftermath of the earthquake.

In line with this argument, [Larrañaga and Herrera \(2010a\)](#) remark that in the regions of Biobio and Maule, the two regions most affected by the earthquake, the 37% and the 22% of the population affected by the earthquake (93% of its population) organized collectively to overcome the damage caused by the earthquake and the provision of security was the main reported reason for collective organization in Biobio and the second (after the provision of water and food) in Maule. Also, [Grandón et al. \(2014\)](#) provide qualitative evidence from the city of Concepción that neighbours cooperated to provide informal guardianship and protect their communities from looting during the week that followed the earthquake.

To explore whether the lasting drop in the incidence of crime in earthquake affected areas was simply driven by an increment in the perceived risk of crime in these municipalities with lasting consequences in terms of adoption of crime prevention measures, I examine whether the effect of the earthquake was significantly different in those treatment municipalities that experienced looting events in the aftermath of the earthquake. For this, I divide the municipalities affected by the earthquake in two separate groups. The first group includes those municipalities affected by the earthquake that experienced looting in the aftermath of the earthquake. The second group includes those municipalities affected by the earthquake that did not. Arguably, the perception of crime in the aftermath of the earthquake was higher among the first treatment group of municipalities. The results are presented in column 9 of table 3.8 in appendix 3.A and show that the magnitude of the effect of the earthquake in areas close to the hypocentre that experienced looting and that did not experience it relative to control municipalities was very similar and the difference between these two magnitudes is statistically indistinguishable from 0 at conventional confidence levels. The latter results suggest that although in the first instance the rise in the perceived risk of crime in earthquake affected areas could have driven the adoption of crime-prevention measures, the rise in the perceived risk of crime in the aftermath of the earthquake cannot explain the observed persistent reduction in the incidence of property crime after the earthquake. Nonetheless, the eruption of looting across some of the municipalities affected by the earthquake could have not been random even among municipalities exposed to the same earthquake intensity and therefore, the results of this analysis should only be interpreted as suggestive.

Table 3.6: Effects of the earthquake on social capital and the adoption of individual and community-based measures to prevent crime

Adoption of crime prevention measures	(1) Share number with neigh. (0/1)	(2) Community vigilance (0/1)	(3) Coord. with local author. (0/1)	(4) Community hires private vigil. (0/1)	(5) Deg (0/1)	(6) Bars in windows or doors (0/1)	(7) Safety lock (0/1)	(8) Alarm (0/1)
<i>Pooled effects</i>								
Earthquake \times Post	0.061* (0.031)	0.062** (0.026)	-0.026 (0.059)	0.009 (0.015)	-0.001 (0.026)	-0.048 (0.038)	-0.064 (0.054)	0.050** (0.022)
Pre-earthq. trends								
F-test: Lead variables $H_0: \beta_{T=-q} = \dots = \beta_{T=-1} = 0$	1.794	1.498	1.495	0.326	1.778	1.475	0.960	6.917***
Observations	67,546	67,546	67,546	67,546	67,546	67,546	67,546	67,546
Dep var. (treatment areas)	0.243	0.098	0.323	0.061	0.405	0.454	0.268	0.065
Social capital: Community assoc.								
	(9) Mothers assoc. (per 100 inhab)	(10) Elderly assoc. (per 100 inhab)	(11) Sport clubs (per 100 inhab)					
<i>Pooled effects</i>								
Earthquake \times Post	0.016* (0.008)	0.005 (0.005)	0.024* (0.013)					
Pre-earthq. trends (F-test)								
F-test: Lead variables $H_0: \beta_{T=-q} = \dots = \beta_{T=-1} = 0$	2.308	0.085	0.063					
Observations	1,033	1,037	1,037					
Dep var. (treatment areas)	0.065	0.095	0.233					

Note: Columns 1-8 estimate at the household level the effect of the earthquake on the adoption of different individual and community-based measures to prevent crime using the pooled effect difference in difference model (equation 3.4). Columns 9-11 estimate at the municipality level the effect of the earthquake on social capital variables using the pooled effect difference in difference model (equation 3.6) and the sample of municipalities for which this information is available. Social capital variables are number of associations per 100 inhabitants, using the population of the municipality in 2009. The models used in these regressions measure the average effect of the earthquake over the post-earthquake period of interest. The effect of interest is yielded by an interaction between the dummy variables that capture whether the municipality is affected by the earthquake and whether the year is after the earthquake. A test for the common trends assumption is reported for every estimation. For this test, I estimate a leads and lags model and use an F-test to examine the joint significance of the lead variables. The mean of the dependent variable is provided for the treatment areas in 2009. Standard errors clustered at the municipality level. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

In order to investigate the relevance of some of the alternative mechanisms, I first assess at the municipality level the short-term effects of exposure to the earthquake on different socioeconomic outcomes that the literature has linked to crime. I estimate the short-term effects of the earthquake on the number of policemen per 100,000 inhabitants, population, poverty rate, extreme poverty rate, unemployment, rate of men 15-29 years old, two measures of income polarization, enrolment in education for individuals 13-25 and municipality budget at the municipality level. All of these factors have been discussed in the literature as potential causes of crime¹⁷. The dependent variable in these regressions is the change in the variable of interest between the first year for which data are available after the earthquake (e.g. 2010 for administrative data and 2011 for variables constructed using the CASEN survey) and the last year before the earthquake (2009 for all variables). The regressions include as control variables the population of the municipality in the year 2009 and the level of the variable of interest in the year 2009.

The estimates are reported in table 3.7 and suggest that proximity to the hypocentre decreased the population of the municipality and increased its unemployment level, poverty and extreme poverty rate. On the other hand, the analysis shows negligible and statistically insignificant effects of earthquake exposure on inequality, number of policemen, budget of the municipality, education enrolment and the rate of men 15-29 years old. Interestingly, table 3.10 in appendix 3.A remarks that, with the exception of extreme poverty, the earthquake did not affect in the short-term any of the variables analysed in the excluded intermediate municipalities. This result is somehow expected because low earthquake intensities are unlikely to damage constructions other than the poorest dwellings that might be more vulnerable.

The results reported in table 3.7 suggest therefore that the lasting drop in property crime rates was not caused by an increase in the presence of policemen or by reconstruction programmes in catastrophic areas reducing unemployment, which has been assessed as a key determinant of crime (Chalfin and McCrary, 2017). Another mechanism that may have contributed to the reduction in the incidence of property crime observed in earthquake affected areas would be a larger incarceration rate in these municipalities. To cope with looting in the aftermath of the earthquake, the Chilean government declared a curfew and deployed the army in the areas affected by riots. If these institutional efforts

¹⁷The evidence on the relevance of most of these factors as drivers of property crime is reviewed in Soares (2004).

led to larger apprehension and incarceration rates, the incidence of crime in earthquake affected municipalities could have dropped as a consequence. However, in the previous section, I show that the earthquake did not increase apprehension rates in the aftermath of the earthquake. Consistently, the results reported in column 9 of table 3.8 in appendix 3.A show that the effect of the earthquake on the incidence of home burglary was not significantly different in treatment municipalities that experienced looting episodes and in those that did not. These two results suggest that the drop in property crime rates in earthquake affected areas was not driven by higher incarceration rates in the aftermath of the earthquake. Furthermore, the lack of a differential effect in municipalities that experienced looting events also indicates that the presence of the army and the curfew, that affected mainly those municipalities that experienced the larger looting events, did not generate any differential effect on the incidence of property crime across municipalities affected by the earthquake. This result dismisses the hypothesis that through temporarily increasing the cost of crime and keeping out of crime some individuals, the curfew and the deployment of the army could have driven the lasting reduction in the incidence of property crime in the municipalities affected by the earthquake.

Another mechanism for lower property crime rates after natural disasters is a reduction in the benefits of crime. Through destroying economic assets and expanding poverty, the earthquake may have decreased the economic returns to some property crimes. In other words, through increasing poverty and destroying assets, the earthquake may have decreased the expected benefit of larceny, robbery or home burglary. Although this argument seems intuitive, theoretical models of economics of crime predict an ambiguous effect of poverty on property crime: Although poverty decreases the economic returns to property crime, it also reduces its opportunity costs. Indeed, the existing empirical evidence shows that different economic shocks increasing poverty in India, Mozambique and Russia have boosted property crime rather than decreasing it (Fafchamps and Minten, 2006; Iyer and Topalova, 2014; Ivaschenko et al., 2012).

Table 3.7: The effects of the earthquake on other sociodemographic and economic variables

	Δ Ln Munic. p/c budget	Δ Ln population	Δ Ln Polic. 100M inhab	Δ Poverty rate	Δ Extreme pov. rate	Δ Unemp. rate	Δ Polariz (75%vs25%)	Δ Polariz. (90%vs10%)	Δ Rate men 15-29	Δ Attending educ. (13-25)
Earthquake municip.	-0.004 (0.014)	-0.005** (0.002)	0.029 (0.023)	0.027** (0.011)	0.015*** (0.004)	0.019** (0.008)	-0.459 (0.409)	-0.080 (1.761)	-0.001 (0.004)	0.014 (0.016)
Observations	157	161	161	140	140	140	140	140	140	140
R-squared	0.005	0.064	0.047	0.158	0.507	0.175	0.433	0.426	0.226	0.275

Note: The table reports the short-term effects of the earthquake on different factors that have been identified in the literature as potential causes of crime. The model estimated is $\Delta Y_i = \beta_0 + \beta_1 \text{Earthquake}_i + \beta_2 Y_{2009_i} + \beta_3 \text{LnPopulation}_{2009_i} + \mu$ where the dependent variable (ΔY) is the change in the variable of interest between the closest available point after the earthquake and the closest available point before the earthquake. Because the data on the budget is at the start of the year, the first relevant post-earthquake year is 2011 for this variable. The first post-earthquake year for which information is available is 2011 for poverty, unemployment, income polarization, age composition and education enrolment, and 2010 for population and policemen. The last pre-earthquake year is 2009 for all the variables. The regressions include as control variables the Ln of population (*LnPopulation*2009) and the variable of interest (*Y*2009) in 2009. The estimation is conducted at the municipality level using OLS and excluding from the estimation the municipalities exposed to a predicted earthquake intensity $5.75 \leq MMI < 7.5$. The difference in the number of observations across the different regressions is explained by the fact that the survey used to construct the poverty, unemployment, polarization, demography and education variables is not implemented in all the Chilean municipalities and the municipality budget data does not include information for all the municipalities. Robust standard errors in parentheses. ***p<0.01; **p<0.05; *p<0.1.

An alternative hypothesis that would help to explain why areas affected by the earthquake experienced strong decreases in crime rates is larger public investments in programmes that may reduce crime in the short- and long-term. In table 3.7 I show that despite the existence of specific transfers from the central government to the municipalities affected by the earthquake (accounting in average for approximately the 3% of the budget of treatment municipalities), exposure to the earthquake did not increase the total municipality budget per inhabitant. However, it is also possible that many of these large investments conducted in damaged areas were not funded by the municipality but directly by the central government. Unfortunately, I do not have the necessary information to test this hypothesis and therefore, I cannot reject the possibility that the decrease in crime was partially explained by a redistribution of public investments and infrastructure towards the areas affected by the earthquake. Nonetheless, we know that if it existed, this effect did not operate through reducing unemployment.

3.10 Conclusions

This study exploits across space variation in exposure to an 8.8 Richter magnitude earthquake in Chile to provide the first evidence on the lasting effects of natural disasters on property crime. For this purpose, property crime data from household victimization surveys and from police records are analysed using a difference in difference strategy. The estimates show that exposure to a very strong earthquake intensity decreased significantly the incidence of home burglary the year of the earthquake. Furthermore, the effect remained constant over the 4 post-earthquake years studied. The results are robust to the use of different sources of data, types of property crime, samples and alternative definitions of treatment and control municipalities. Although I cannot rule out the possibility that these results are affected by indirect effects of the earthquake in control municipalities, the sharp break in the crime trend in treatment municipalities the year of the earthquake and the smooth trend in control municipalities the same year suggest that if existent, such an indirect effect would be small and could not explain entirely the results.

The study also explores some of the mechanisms through which the earthquake may have reduced property crime in the medium and long-term. An important driver of this effect was the lasting boost in the adoption of community-based measures to prevent

crime in earthquake affected areas. More broadly, the results are consistent with the stream of the literature that argues that natural disasters increase the level of cooperation within neighbourhoods and the strength of community life leading to larger levels informal guardianship in affected communities and increasing the cost of committing crime after catastrophic events. Furthermore, the evidence highlights the role played by social capital and cooperation at the community level in reducing crime, a question that has not been empirically investigated so far. Alternative mechanisms to explain the lasting drop in the incidence of property crime after the earthquake such as an increase in the number of policemen in areas affected by the earthquake, higher incarceration rates, crime displacement, an increase in the perceived risk of crime, lasting effects of the curfew and army deployment and an increase in employment due to the reconstruction programmes are tested and ruled out in the light of the results. However, natural disasters are complex phenomena with numerous consequences and therefore, I cannot dismiss the possibility that the lasting drop in the prevalence of property crime after the earthquake was also channelled through other mechanisms not examined in this study.

Appendix 3.A Robustness Checks: Using All Observations and Heterogeneity of Effects

Table 3.8: Effects of the earthquake on home burglary: Different samples, municipality time trends and heterogeneity of effects

Different samples	(1) Home burglary (0/1)	(2) Home burglary (0/1)	(3) Home burglary (0/1)	(4) Home burglary (0/1)	(5) Home burglary (0/1)	(6) Home burglary (0/1)	(7) Home burglary (0/1)	(8) Home burglary (0/1)
<i>Pooled effects</i>								
Earthquake \times Post	-0.021*** (0.005)	-0.021*** (0.005)	-0.021*** (0.005)	-0.019*** (0.005)	-0.023** (0.010)	-0.023** (0.010)	-0.023** (0.010)	-0.026** (0.011)
Intermediate areas \times Post		-0.014*** (0.004)	-0.015*** (0.005)			-0.011 (0.009)	-0.010 (0.010)	
Pre-earthq. trends								
F-test: Lead variables								
$H_0: \beta_{\tau=-q} = \dots = \beta_{\tau=-1} = 0$	0.238	0.242	0.241	0.174	1.554	1.583	1.572	1.094
Municip. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municip. time trends	No	No	No	No	Yes	Yes	Yes	Yes
Observations	67,540	177,889	111,622	57,100	67,540	177,889	111,622	57,100
Treatment areas	MMI ≥ 7.5	MMI ≥ 7.5	MMI ≥ 7.5	MMI ≥ 7.5	MMI ≥ 7.5	MMI ≥ 7.5	MMI ≥ 7.5	MMI ≥ 7.5
Control areas	MMI < 5.75	MMI < 5.75	MMI < 5.75	MMI < 5.75	MMI < 5.75	MMI < 5.75	MMI < 5.75	MMI < 5.75
Intermediate areas	Excluded	5.75 \leq MMI < 7.5	5.75 \leq MMI < 7.5 Santiago excluded	Excluded	Excluded	5.75 \leq MMI < 7.5	5.75 \leq MMI < 7.5 Santiago excluded	Excluded
Tsunami affected municip	Included	Included	Included	Excluded	Included	Included	Included	Excluded
Heterog. of effects	(9) Home burglary (0/1)	(10) Home burglary (0/1)						
<i>Pooled effects</i>								
Earthquake \times Post	-0.023*** (0.005)							
(Munic with looting)								
Earthquake \times Post	-0.019*** (0.006)							
(Munic without looting)								
Earthquake \times Post		-0.024*** (0.005)						
(Munic δ CBS=1)								
Earthquake \times Post		-0.011*** (0.004)						
(Munic δ CBS=0)								
Observations	67,540	67,540						
Treatment areas	MMI ≥ 7.5	MMI ≥ 7.5						
Control areas	MMI < 5.75	MMI < 5.75						

Note: Columns 1-8 examine the pooled effects of the earthquake on home burglary over the period of interest (equation 3.4) using different samples and specifications. The effect of interest is captured by an interaction between the dummy variables that capture whether the municipality is affected by the earthquake and whether the year is after the earthquake. A test for the common trends assumption is reported for every estimation. For this test, I estimate a leads and lags model and use an F-test to examine the joint significance of the lead variables. Columns 9-10 estimate the pooled effect of the earthquake using the same control group and splitting the treatment municipalities in two different groups: Those treatment municipalities that experienced looting (column 9) or an increase in the provision of community-based crime prevention measures (column 10) and those that did not. All the regressions are estimated at the household level using ENUSC data. Standard errors clustered at the municipality level. ***p<0.01; **p<0.05; *p<0.1.

Table 3.9: Impact estimates (OLS): Short-term effects of the earthquake on different types of property crimes and on individuals apprehended (SPD data)

	Δ Home burglary (per 1,000 inhab)	Δ Larceny (per 1,000 inhab)	Δ Non-home burglary (per 1,000 inhab)	Δ Motor-vehicle thefts (per 1,000 inhab)	Δ Robbery (per 1,000 inhab)	Δ Apprehended (per 1,000 inhab)
<i>Sample A: March 2010 - Jan 2010</i>						
Earthquake municip.	-0.104*** (0.028)	-0.194*** (0.051)	0.042 (0.050)	-0.003 (0.007)	-0.021 (0.013)	-0.151** (0.067)
Intermediate municip.	-0.054* (0.029)	-0.094* (0.052)	-0.033 (0.027)	-0.007 (0.008)	0.050** (0.022)	-0.089 (0.065)
<i>Sample B: Feb 2010 - Jan 2010</i>						
Earthquake municip.	-0.074 (0.047)	-0.072 (0.055)	-0.068 (0.066)	-0.016* (0.009)	0.031 (0.022)	0.074 (0.074)
Intermediate municip.	-0.054 (0.043)	0.016 (0.053)	-0.043 (0.047)	-0.013 (0.010)	0.088*** (0.031)	0.034 (0.065)
Observations	345	345	345	345	345	345
Av. rate Jan 2010 (Treat mun)	0.328	0.527	0.215	0.057	0.138	0.542

Note: The regressions estimated use monthly data from police records (SPD database) and OLS methods to estimate at the municipality level the short term effects of the earthquake on property crime and on individuals apprehended. The equation estimated is $\Delta Y = \beta_0 + \beta_1 \text{Earthquake} + \beta_2 Y + \mu$ where the dependent variable ΔY is the difference in crime rates/individuals apprehended between March 2010 (the first month after the earthquake) and January 2010 (the last month before the earthquake) in sample A and the difference in crime rates/individuals apprehended between February 2010 (the month of the earthquake) and January 2010 in sample B. Y measures the crime rate/number of people apprehended in January 2010. Municipalities exposed to a predicted $5.75 \leq MMI < 7.5$ are included in the sample as a separate treatment group (intermediate exposure). Robust standard errors in parentheses. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table 3.10: The effects of the earthquake on other sociodemographic and economic variables

	Δ Ln Munic. p/c budget	Δ Ln population	Δ Ln Polic. 100M inhab	Δ Poverty rate	Δ Extreme pov. rate	Δ Unemp. rate	Δ Polariz (75%vs25%)	Δ Polariz. (90%vs10%)	Δ Rate men 15-29	Δ Attending educ. (13-25)
Earthquake municip.	-0.006 (0.014)	-0.005* (0.002)	0.023 (0.023)	0.021** (0.010)	0.014*** (0.005)	0.017** (0.008)	-0.427 (0.419)	-0.143 (1.734)	-0.002 (0.004)	0.016 (0.016)
Intermediate municip.	-0.001 (0.014)	-0.001 (0.002)	0.017 (0.021)	0.001 (0.006)	0.008** (0.003)	0.005 (0.005)	-0.082 (0.401)	-0.107 (1.281)	-0.003 (0.003)	0.015 (0.013)
Observations	340	345	345	324	324	324	324	324	324	324
R-squared	0.012	0.045	0.020	0.151	0.473	0.169	0.202	0.429	0.217	0.346

Note: The table reports the short-term effects of the earthquake on different factors that have been identified in the literature as potential causes of crime. The model estimated is $\Delta Y_i = \beta_0 + \beta_1 \text{Earthquake}_i + \beta_2 Y2009_i + \beta_3 \text{LnPopulat}2009_i + \mu$ where the dependent variable (ΔY) is the change in the variable of interest between the closest available point after the earthquake and the closest available point before the earthquake. Because the data on the budget is at the start of the year, the first relevant post-earthquake year is 2011 for this variable. The first post-earthquake year for which information is available is 2011 for poverty, unemployment, income polarization, age composition and education enrolment, and 2010 for population and policemen. The last pre-earthquake year is 2009 for all the variables. The regressions include as control variables the Ln of population ($\text{LnPopulat}2009$) and the variable of interest ($Y2009$) in 2009. The estimation is conducted at the municipality level using OLS. Municipalities exposed to a predicted earthquake intensity $5.75 \leq MMI < 7.5$ are included as a separate treatment group (intermediate municipalities). The difference in the number of observations across the different regressions is explained by the fact that the survey used to construct the poverty, unemployment, polarization, demography and education variables is not implemented in all the Chilean municipalities and the municipality budget data does not include information for all the municipalities. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1.

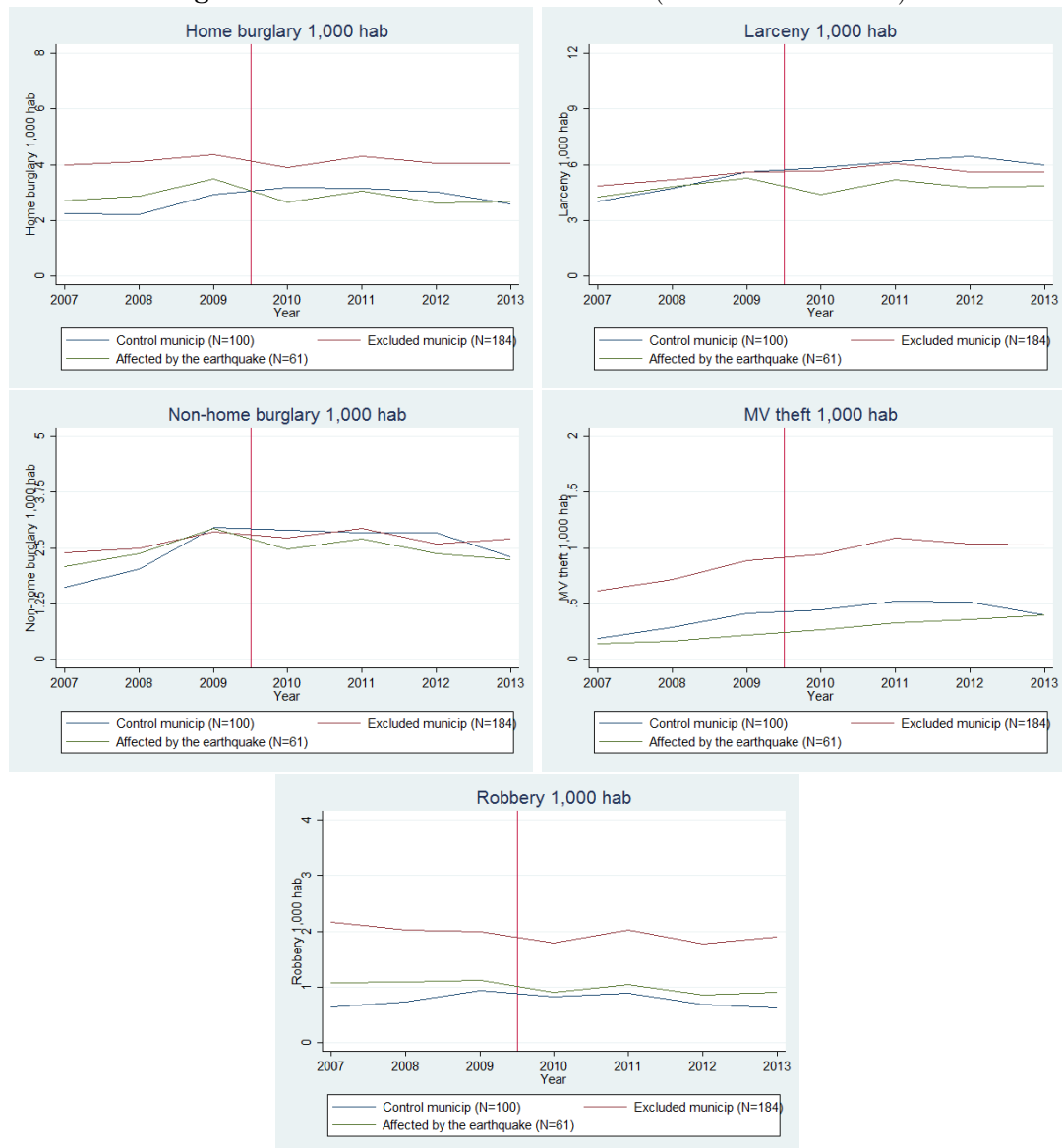
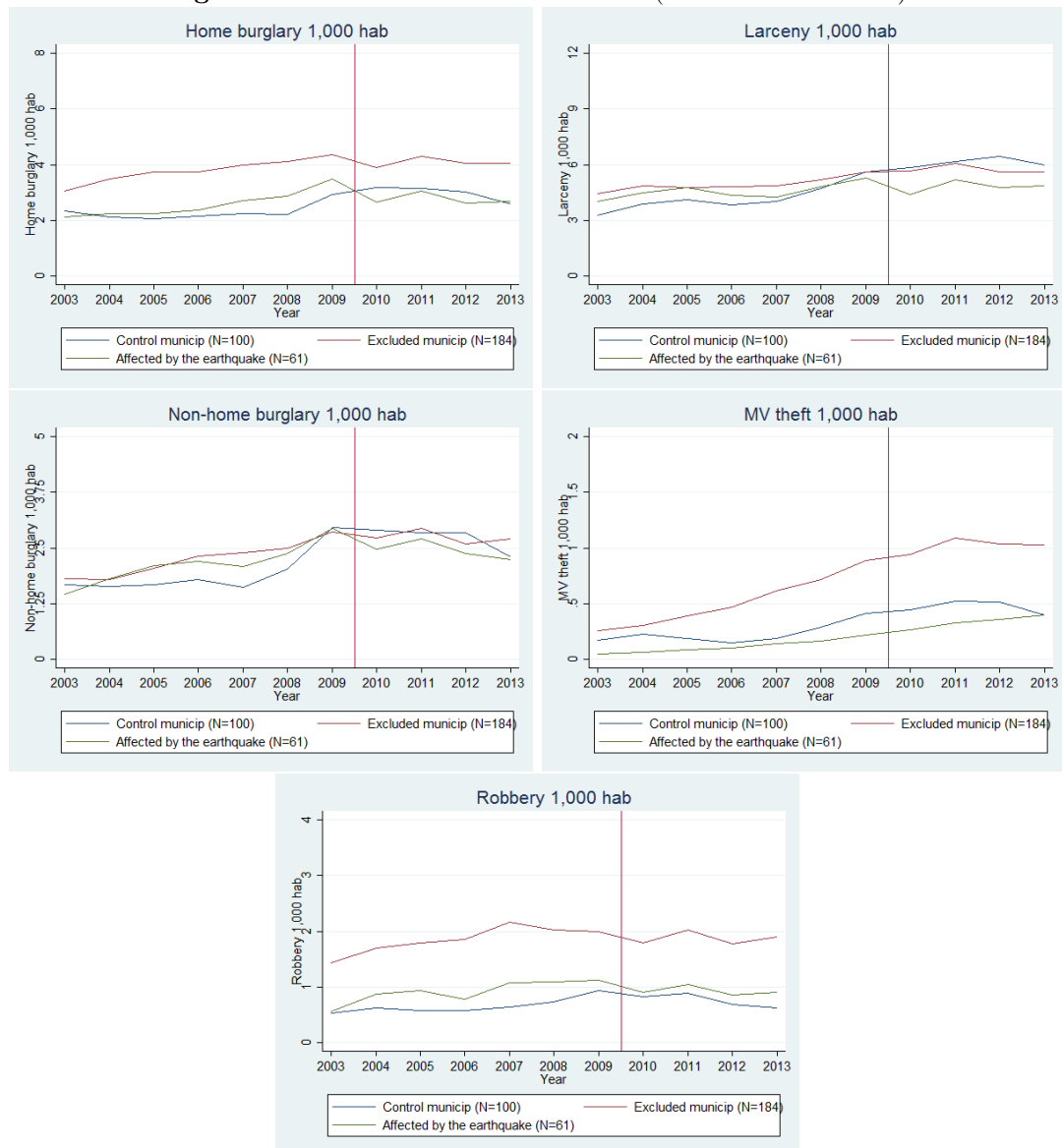
Figure 3.6: Incidence of crime over time (SPD data 2007-2013)

Figure 3.7: Incidence of crime over time (SPD data 2003-2013)

Appendix 3.B Earthquake Intensity Scales

Modified Mercalli intensity (MMI) scale measures the destruction capacity of an earthquake rather than the current destruction that it generates. Given a Richter magnitude, the Modified Mercalli scale in a place depends on the distance to the hypocentre and on the topography of the place. The interpretation of some of the values of the Modified Mercalli scale relevant for this study is reported below.

- MMI IX (Violent): Damage considerable in specially designed structures; well-designed frame structures thrown out of plumb. Damage great in substantial buildings, with partial collapse. Buildings shifted off foundations.
- MMI VIII (Severe): Damage slight in specially designed structures; considerable damage in ordinary substantial buildings with partial collapse. Damage great in poorly built structures. Fall of chimneys, factory stacks, columns, monuments walls. Heavy furniture overturned.
- MMI VII (Very strong): Damage negligible in buildings of good design and construction; slight to moderate in well-built ordinary structures; considerable damage in poorly built or badly designed structures; some chimneys broken.
- MMI VI (Strong): Felt by all, many frightened. Some heavy furniture moved; a few instances of fallen plaster. Damage slight.
- MMI V Moderate: Felt by nearly everyone; many awakened. Some dishes, windows broken. Unstable objects overturned. Pendulum clocks may stop.

Medvedev-Sponheuer-Karnik (MSK) scale measures the severity of ground shaking on the basis of observed effects in an area affected by an earthquake. Given a Richter magnitude, the MSK scale in a place depends on the distance to the hypocentre and on the topography of the place. The interpretation of some of the values of the MSK scale relevant for this study is reported below.

- MSK IX Destructive: General panic. People may be forcibly thrown to the ground. Waves are seen on soft ground. Substandard structures collapse. Substantial damage to well-constructed structures. Underground pipelines ruptured. Ground fracturing, widespread landslides.
- MSK VIII Damaging: Many people find it difficult to stand, even outdoors. Furniture may be overturned. Waves may be seen on very soft ground. Older structures partially collapse or sustain considerable damage. Large cracks and fissures opening up, rockfalls.
- MSK VII Very strong: Most people are frightened and try to run outdoors. Furniture is shifted and may be overturned. Objects fall from shelves. Water splashes from containers. Serious damage to older buildings, masonry chimneys collapse. Small landslides.
- MSK VI Strong: Felt by most indoors and by many outdoors. A few persons lose their balance. Many people are frightened and run outdoors. Small objects may fall and furniture may be shifted. Dishes and glassware may break. Farm animals may be frightened. Visible damage to masonry structures, cracks in plaster. Isolated cracks on the ground.
- MSK V Fairly strong: Felt indoors by most, outdoors by few. A few people are frightened and run outdoors. Many sleeping people awake. Observers feel a strong shaking or rocking of the whole building, room or furniture. Hanging objects swing considerably. China and glasses clatter together. Doors and windows swing open or shut. In a few cases window panes break. Liquids oscillate and may spill from fully filled containers. Animals indoors may become uneasy. Slight damage to a few poorly constructed buildings

Appendix 3.C Maps: Treatment, Control and Excluded Municipalities under the Use of Different Distance Thresholds

Figure 3.8: Treatment and Control areas

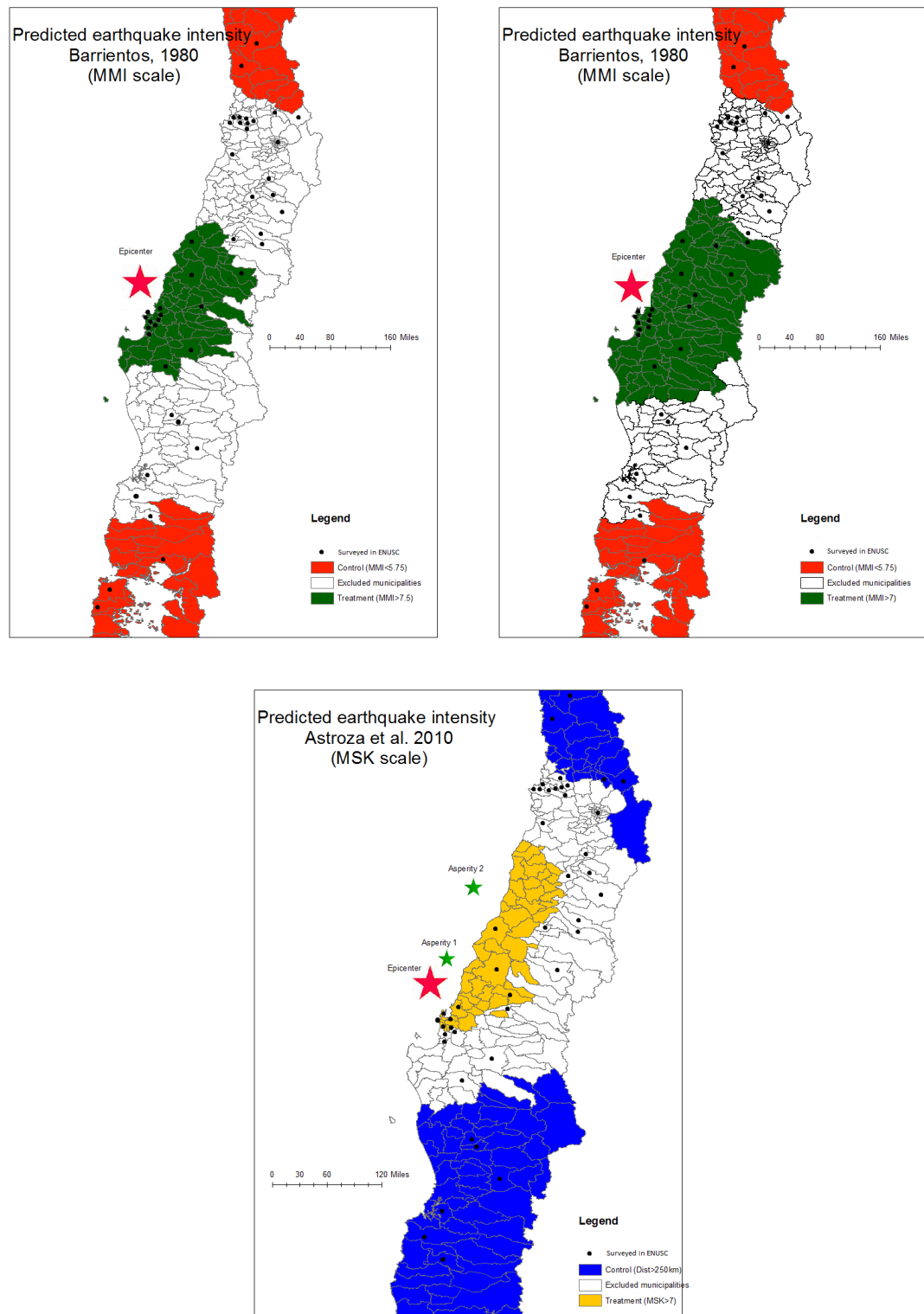
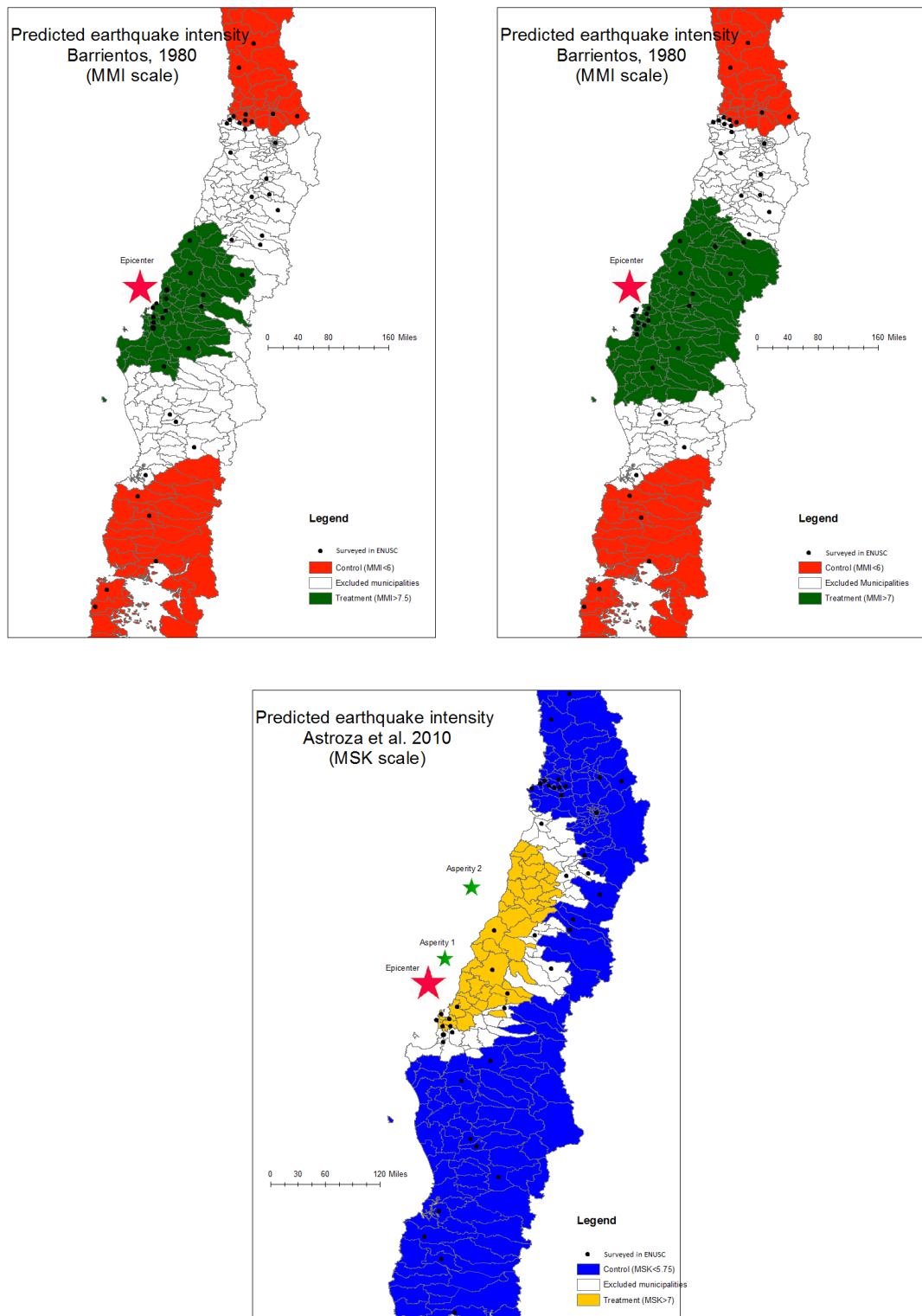


Figure 3.9: Treatment and Control areas

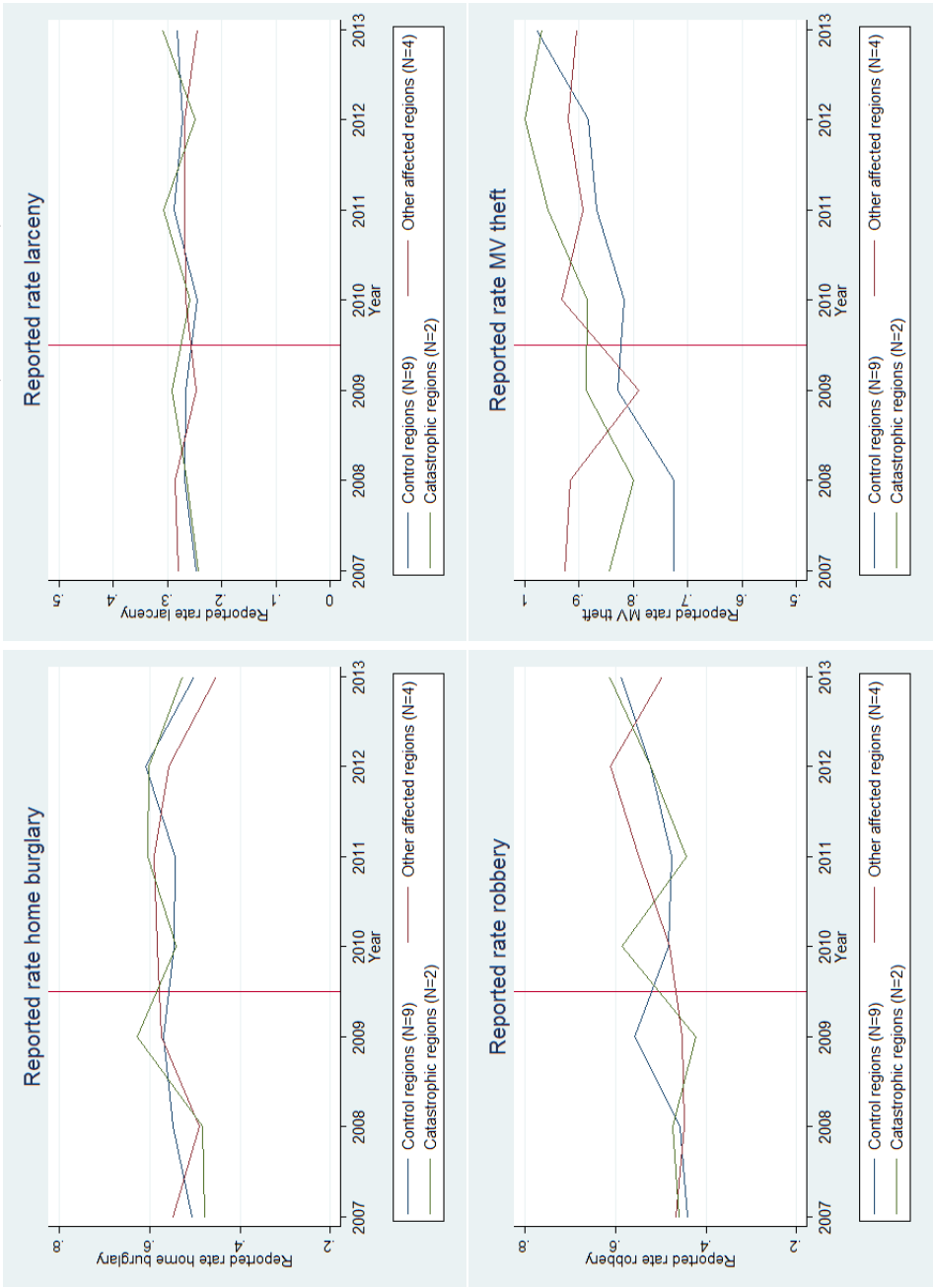
Appendix 3.D Reporting Rate for Different Types of Crime (ENUSC Data): Analysis at the Regional Level

Table 3.11: Effect of the earthquake on the probability of reporting a crime to the police and mean reporting rates (regional level analysis)

	Share crime reported to the police
POST \times Catastrophic regions	0.021 (0.020)
POST \times Other affected regions	-0.014 (0.019)
Type of crime fixed effect	Yes
N Observations	412
R2	0.688
Type of crime	Share reported to the police
Home burglary	0.546
Larceny	0.268
Motor vehicle theft	0.862
Robbery	0.504

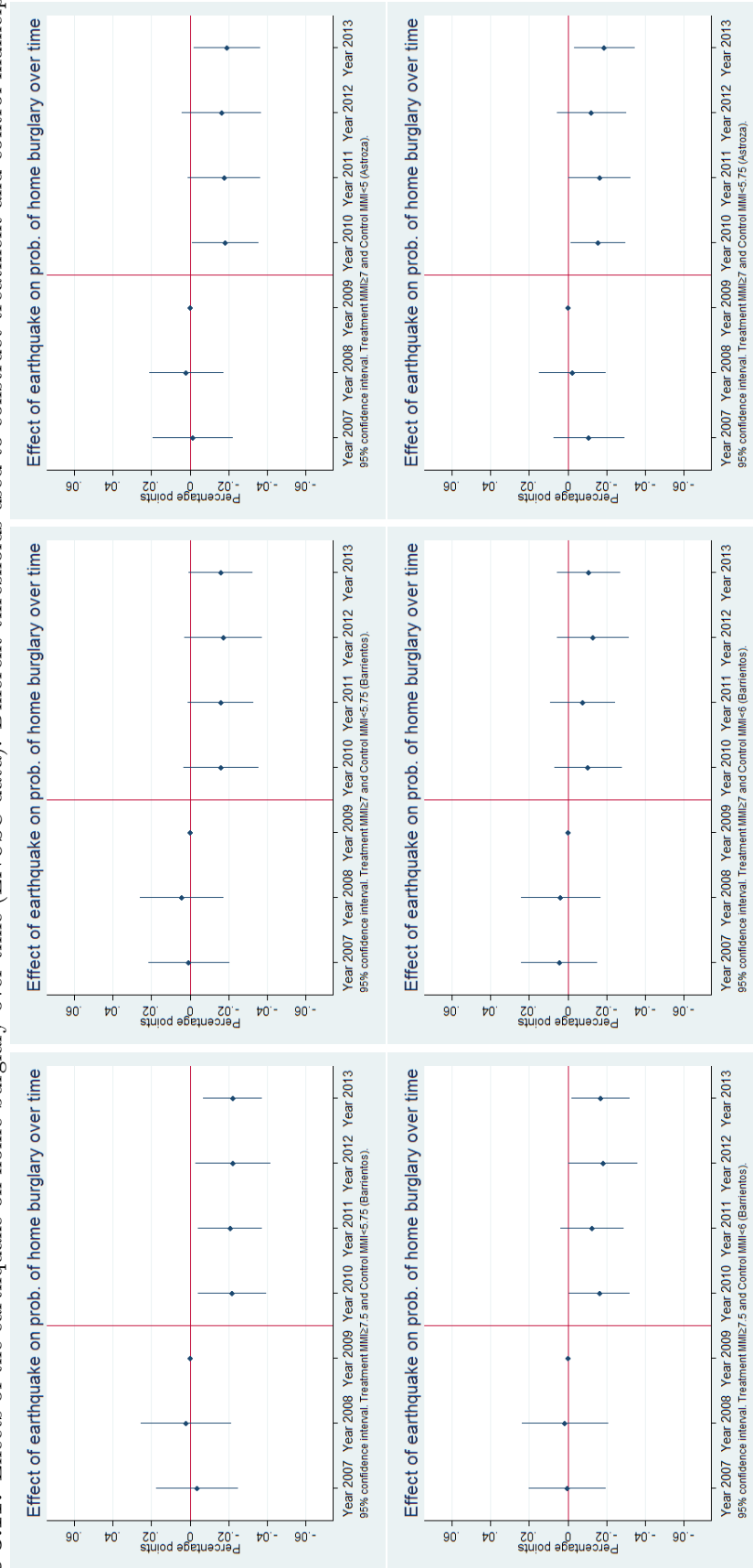
Note: The control regions are Tarapaca, Antofagasta, Arica y Parinacota, Coquimbo, Atacama, Los Rios, Los Lagos, Aysen, Magallanes. Information on reported crime for the regions of Los Rios and Arica y Parinacota for the years 2007 and 2008 is not available. Catastrophic regions include the regions of Maule and Biobio and other affected regions include the regions of Santiago, Valparaiso, Araucania and Libertador O'Higgins. The regressions also include a dummy variable that is equal to 1 for the years after the earthquake and a vector of exposure to the earthquake fixed effects (catastrophic, other affected regions and control). The dependent variable in the regression is the share of crime reported to the police in each region and for each type of property crime (larceny, motor-vehicle theft, robbery and home burglary). Robust standard errors in parentheses. ***p<0.01; **p<0.05; *p<0.1

Figure 3.10: Evolution of reporting rate by type of crime (ENUSC data)



Appendix 3.E Additional Graphs

Figure 3.11: Effects of the earthquake on home burglary over time (ENUSC data): Different thresholds used to construct treatment and control municipalities



Conclusions

Using econometric methods traditionally applied for the identification of causal effects and policy evaluation, this thesis provides empirical evidence on three central issues for the well-being of disadvantaged people in developing countries.

Chapter 1 uses a legal change in Ethiopia as a case study to show the beneficial effects on infant mortality rates of (a) increasing the legal age of marriage for women and of (b) delaying women's age at cohabitation during teenage years. The size of the effect of a one-year delay in women's age at cohabitation during teenage years on the probability of infant mortality of the first born child is comparable to the joint effect on child mortality at the village level of moving from 0% coverage of measles, BCG, DPT, Polio and Maternal Tetanus vaccinations to 100%. The analysis of mechanisms reveals that the reduction in infant mortality is strongly linked to the effect of delaying cohabitation on the age of women at first birth.

The contribution of chapter 1 is threefold. First, it provides the first empirical evidence on the socioeconomic effects of laws raising the legal age of marriage for women. Second, it assesses for the first time the causal effect of early cohabitation on infant mortality. Third, the chapter uses a RDD approach that can be used in Ethiopia or in other countries such as Albania or Jordan that have approved similar laws over the last decades to expand the analysis on the effects of child marriage to other key outcomes and settings, and to examine the robustness of the results from studies relying on correlation analysis or instrumental variables.

Future research using comprehensive data on postnatal and antenatal behaviours could also disentangle whether the *age at first birth* mechanism is in itself driven by purely biological reasons or by a larger adoption of adequate antenatal and postnatal health practices by women that are older at first birth. Additional research is also needed to understand whether the magnitude of the effect of the age at cohabitation on the infant

mortality of the first born child depends on the gender of the infant. The larger incidence of infant mortality among boys observed in both developed and developing countries seems to be driven by a slower development of the immune system in boys and by a larger probability of pre-term birth when the infant is a boy (Drevenstedt et al., 2008). If biological maturity at first birth is the main mechanism through which age at cohabitation affects infant mortality, one may hypothesise that the impact of early cohabitation on infant mortality could also depend on the gender of the infant.

Using longitudinal data from northern Ghana, chapter 2 assesses whether parents allocate human capital investments reinforcing or correcting cognitive differences between siblings. The analysis reveals that consistent with the predictions of Becker's model for intra-household allocation of resources, parents allocate more schooling to the children of the household that are cognitively more able. Furthermore, the results show that the magnitude of this effect does not seem to depend on household characteristics such as wealth, household size or polygyny status.

In policy terms, the results suggest that since parents may prefer to focus their schooling investments in the more able children rather than spreading them across all of them, supply-side educational interventions such as reducing class sizes or providing more books may not benefit the less able children.

The contribution of this chapter is twofold. On the one hand, it adds to the thin literature that examines empirically the role of cognitive skills in the allocation of schooling across siblings. On the other hand, this is the first study that tests whether the magnitude of the effect of cognitive skills on the allocation of schooling across siblings varies across different types of households.

More research is however needed to determine whether parents allocate more schooling to more able children because these children have larger returns on human capital investments as suggested by Becker or because parents have a preference for these children regardless of their returns. More generally, one way of testing the functioning of Becker's model would be through examining other predictions of the model. Two of them that are worth to be investigated for Ghana are whether the reinforcing mechanism is also found for health endowments and whether less endowed children receive more nonhuman capital transfers.

Chapter 3 uses the 8.8 Richter magnitude earthquake that struck Chile in February

2010 as a case study to investigate the lasting effects of natural disasters on property crime. The results show that municipalities close to the earthquake hypocentre experienced a lasting decrease in the incidence of property crime relative to those municipalities unaffected by the earthquake. Furthermore, the analysis reveals that the reduction in crime was closely linked to the positive effect of the earthquake on the strength of community links and on the adoption of community-based measures to prevent crime.

This chapter is the first empirical study assessing the effects of natural disasters on crime over post-disaster periods of time longer than a year. In addition, the evidence presented in this chapter remarks the key role that community links and informal guardianship can play to prevent crime. Although previous evidence discusses the link between social capital and crime in the Netherlands and in Italy ([Buonanno et al., 2009](#); [Akcomak and Ter Weel, 2012](#)), none of them focuses on community links and the provision of informal guardianship. In this sense, the chapter opens the debate and calls for more research assessing how community links and the provision of informal guardianship can contribute to reducing crime, particularly in settings where formal institutions have limited capacity to enforce the law.

Bibliography

- Addy, A. (2013). Contextualising the Underperformance of Rural Education in Northern Ghana: Management Approach. *International Journal of ICT and Management*, 1(3).
- Adhikari, R. (2003). Early marriage and childbearing: risks and consequences. In Bott, S., Jejeebhoy, S., Shah, I., and Puri, C., editors, *Towards adulthood: exploring the sexual and reproductive health of adolescents in South Asia*, pages 62–66. World Health Organization.
- Akcomak, I. S. and Ter Weel, B. (2012). The impact of social capital on crime: Evidence from the Netherlands. *Regional Science and Urban Economics*, 42(1-2):323–340.
- Akresh, R., Bagby, E., de Walque, D., and Kazianga, H. (2012). Child Ability and Household Human Capital Investment Decisions in Burkina Faso. *Economic Development and Cultural Change*, 61(1):157 – 186.
- Alderman, H. (1995). *Public Schooling Expenditures in Rural Pakistan: Efficiently Targeting Girls and a Lagging Region*. Distributed by ERIC Clearinghouse [Washington, D.C.].
- Allendorf, K. (2007). Do Women’s Land Rights Promote Empowerment and Child Health in Nepal? *World Development*, 35(11):1975–1988.
- Angrist, J. D. and Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist’s Companion* Princeton University Press. Princeton University Press.
- Appleton, S. (2000). Education and health at the household level in sub-Saharan Africa. CID Working Papers 33, Center for International Development at Harvard University.
- Asadullah, M. and Wahhaj, Z. (2016). Early Marriage, Social Networks and the Transmission of Norms. Studies in Economics 1602, School of Economics, University of Kent.

- Asadullah, N., Alim, A., Khatoon, F., and Chaudhury, N. (2016). Maternal Early Marriage and Cognitive Skills Development: An Intergenerational Analysis. Technical report, Unpublished working paper.
- Ashraf, N., Bau, N., Nunn, N., and Voena, A. (2016). Bride Price and Female Education. NBER Working Papers 22417, National Bureau of Economic Research, Inc.
- Astroza, M., Cabezas, F., Moroni, M., Massone, L., Ruiz, S., Parra, E., Cordero, F., and Mottadelli, A. (2010). Intensidad Sísmica en el Area de Daños del Terremoto del 27 de Febrero de 2010. Fcfm, Universidad de Chile.
- Autor, D. H. (2003). Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing. *Journal of Labor Economics*, 21(1):1–42.
- Ayalew, T. (2005). Parental Preference, Heterogeneity, and Human Capital Inequality. *Economic Development and Cultural Change*, 53(2):381–407.
- Bailey, K. (2009). An Evaluation of the Impact of Hurricane Katrina on Crime in New Orleans, Louisiana. Technical report, Unpublished Applied Research Project for a Masters of Public Administration, Department of Political Science, Texas State University.
- Barrientos, S. (1980). Regionalización Sísmica de Chile. Technical report, Department of Geophysics. Universidad de Chile.
- Barrios, Y., Sanchez, S., Nicolaidis, C., Garcia, P., Gelaye, B., Zhong, Q., and Williams, M. (2015). Childhood Abuse and Early Menarche among Peruvian Women. *Journal of Adolescent Health*, 56(2):197–202.
- Becker, G. S. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy*, 76:169.
- Becker, G. S. (1981). *A Treatise on the Family*. Harvard University Press.
- Belasen, A. R. and Polachek, S. W. (2008). How Hurricanes Affect Wages and Employment in Local Labor Markets. *American Economic Review*, 98(2):49–53.
- Bell, B., Jaitman, L., and Machin, S. (2014). Crime Deterrence: Evidence From the London 2011 Riots. *Economic Journal*, (576):480–506.

- Bound, J., Jaeger, D. A., and Baker, R. M. (1995). Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak. *Journal of the American Statistical Association*, 90(430):443–450.
- Briggs, D. C. (2004). Causal Inference and the Heckman Model. *Journal of Educational and Behavioral Statistics*, 29(4):397–420.
- Buonanno, P., Montolio, D., and Vanin, P. (2009). Does Social Capital Reduce Crime? *Journal of Law and Economics*, 52(1):145–170.
- Calonico, S., Cattaneo, M. D., Farrell, M., and Titiunik, R. (2016). Regression Discontinuity Designs Using Covariates. Technical report, University of Michigan.
- Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust Nonparametric Confidence Intervals for Regression Discontinuity Designs. *Econometrica*, 82:2295–2326.
- Campbell, J. (2002). Health Consequences of Intimate Partner Violence. *The Lancet*, 359(9314):1331–1336.
- Card, D. and Krueger, A. B. (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *American Economic Review*, 84(4):772–93.
- Cavallo, E., Galiani, S., Noy, I., and Pantano, J. (2010). Catastrophic Natural Disasters and Economic Growth. Research Department Publications 4671, Inter-American Development Bank, Research Department.
- CEPAL (2010). Terremoto en Chile: Una Primera Mirada. Technical report, CEPAL.
- Chalfin, A. and McCrary, J. (2017). Criminal Deterrence: A Review of the Literature. *Journal of Economic Literature*, 55(1):5–48.
- Chari, A., Heath, R., Maertens, A., and Fatima, F. (2017). The Causal Effect of Maternal Age at Marriage on Child Wellbeing: Evidence from India. *Journal of Development Economics*, 127:42 – 55.
- Chiappori, P. and Meghir, C. (2014). Intrahousehold Inequality. Cowles Foundation Discussion Papers 1948, Cowles Foundation for Research in Economics, Yale University.

- Cohen, L. E. and Felson, M. (1979). Social change and crime rate trends: a routine activity approach. *American Sociological Review*, pages 588–608.
- Contreras, M. and Winckler, P. (2013). Pérdidas de vidas, viviendas, infraestructura y embarcaciones por el tsunami del 27 de Febrero de 2010 en la costa central de Chile. *Obras y Proyectos*, 14:6–19.
- Cromwell, P., Dunham, R. G., Akers, R., and Lanza-Kaduce, L. (1995). Routine Activities and Social Control in the Aftermath of a Natural Catastrophe. *Disaster Prevention and Management*, 3(3):56–69.
- Cueto, S., Leon, J., and Muñoz, I. G. (2014). *Educational Opportunities and Learning Outcomes of Children in Peru: A Longitudinal Model*, pages 245–267. Palgrave Macmillan UK, London.
- Curtis, T., Miller, B. C., and Berry, E. (2000). Changes in Reports and Incidence of Child Abuse following Natural Disasters. *Child Abuse and Neglect*, 24(9):1151 – 1162.
- Datar, A., Kilburn, R., and Loughran, D. (2010). Endowments and Parental Investments in Infancy and Early Childhood. *Demography*, 47(1):145–162.
- Dixon, R. (1971). Cross-Cultural Variations in Age at Marriage and Proportions Never Marrying. *Population Studies*, 25(2):215–233.
- Drevenstedt, G. L., Crimmins, E. M., Vasunilashorn, S., and Finch, C. E. (2008). The rise and fall of excess male infant mortality. *PNAS*, 105(13):5016–5021.
- Duffo, E. and Udry, C. (2004). Intrahousehold Resource Allocation in Cote d’Ivoire: Social Norms, Separate Accounts and Consumption Choices. NBER Working Papers 10498, National Bureau of Economic Research, Inc.
- Dunn, T. A. and Phillips, J. W. (1997). The Timing and Division of Parental Transfers to Children. *Economics Letters*, 54(2):135–137.
- Dynes, R. and Quarantelli, E. (1980). Helping Behaviour in Large Scale Disaster. In *Participation in Social and Activities*, pages 339–354. Jossey-Bass.
- Ehrlich, I. (1973). Participation in Illegitimate Activities: A Theoretical and Empirical Investigation. *Journal of Political Economy*, 81(3):521–65.

- Elborgh-Woytek, K., Newiak, M., and Newiak, K. (2007). New Insights on Preventing Child Marriage: A Global Analysis of Factors and Programs. Technical report.
- Elborgh-Woytek, K., Newiak, M., and Newiak, K. (2013). Women, Work, and the Economy: Macroeconomic Gains from Gender Equity. Technical report.
- Enarson, E., Fothergill, A., and Peek, L. (2006). Gender and disaster: Foundations and directions. In *Handbook of Disaster Research*, pages 130–146. New York: Springer.
- Ermisch, J. and Francesconi, M. (2000). Educational Choice, Families, and Young People’s Earnings. *Journal of Human Resources*, 35(1):143–176.
- Fafchamps, M. and Minten, B. (2006). Crime, Transitory Poverty, and Isolation: Evidence from Madagascar. *Economic Development and Cultural Change*, 54(3):579–603.
- Field, E. and Ambrus, A. (2008). Early Marriage, Age of Menarche, and Female Schooling Attainment in Bangladesh. *Journal of Political Economy*, 116(5):881–930.
- Frailing, K. and Harper, D. (2007). Crime and Hurricanes in New Orleans. In *The Sociology of Katrina: Perspectives on a Modern Catastrophe*, pages 51–68. New York: Springer.
- Friesema, P., Caporaso, J., Goldstein, G., Lineberry, R., and McCleary, R. (1979). *Aftermath: Communities after Natural Disasters*. Beverly Hills, CA: Sage Publications.
- Frijters, P., Johnston, D. W., Shah, M., and Shields, M. A. (2010). Intra-household Resource Allocation: Do Parents Reduce or Reinforce Child Cognitive Ability Gaps? IZA Discussion Papers 5153, Institute for the Study of Labor (IZA).
- Glewwe, P. (1999). *The Economics of School Quality Investments in Developing Countries: An Empirical Study of Ghana*. Palgrave Macmillan.
- Goody, J. (1990). *The Oriental, the Ancient and the Primitive: Systems of Marriage and the Family in the Pre-Industrial Societies of Eurasia*. Cambridge University Press.
- Grandón, P., Acuna, S., Briese, C., Chovar, P., Hernández, A., and Orellana, F. (2014). Saqueos y Autodefensa. Impacto Social en Chile Post Terremoto. *Ajayu*, 12(2):187–206.
- Griliches, Z. (1979). Sibling Models and Data in Economics: Beginnings of a Survey. *Journal of Political Economy*, 87(5):S37–64.

- GSS (2015). Ghana Demographic and Health Survey 2014. Technical report, Ghana Statistical Service, Global Health Service and The DHS Programme.
- Hallward-Driemeier, M. and Gajigo, O. (2015). Strengthening Economic Rights and Women's Occupational Choice: The Impact of Reforming Ethiopia's Family Law. *World Development*, 70:260–273.
- Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, 47(1):153–161.
- Hicks, J. and Hicks, D. (2015). Lucky Late Bloomers? The Consequences of Early Marriage for Women in Western Kenya. Technical report, Unpublished working paper.
- Hochguertel, S. and Ohlsson, H. (2009). Compensatory inter vivos gifts. *Journal of Applied Econometrics*, 24(6):993–1023.
- Hume, D. (2000). *Treatise of Human Nature*. Oxford: Oxford University Press.
- Ivaschenko, O., Nivorozhkin, A., and Nivorozhkin, E. (2012). The Role of Economic Crisis and Social Spending in Explaining Crime in Russia. *Eastern European Economics*, 50(4):21–41.
- Iyer, L. and Topalova, P. B. (2014). Poverty and Crime: Evidence from Rainfall and Trade Shocks in India. Technical report, Harvard Business School BGIE Unit Working Paper No. 14-067.
- Jensen, R. and Thornton, R. (2003). Early Female Marriage in the Developing World. *Gender and Development*, 11(2):9–19.
- Jere R. Behrman, Mark R. Rosenzweig, P. T. (1994). Endowments and the Allocation of Schooling in the Family and in the Marriage Market: The Twins Experiment. *Journal of Political Economy*, 102(6):1131–1174.
- Jere R. Behrman, Robert A. Pollak, P. T. (1982). Parental Preferences and Provision for Progeny. *Journal of Political Economy*, 90(1):52–73.
- Jones, N., Tefera, B., Emirie, G., Gebre, B., Berhanu, K., Presler-Marshall, E., Walker, D., Gupta, T., and Plank, G. (2016). One Size Does Not Fit All: The Patterning and

- Drivers of Child Marriage in Ethiopia's Hotspot Districts. Technical report, UNICEF and ODI.
- Karapanou, O. and Papadimitriou, A. (2010). Determinants of Menarche. *Reproductive Biology and Endocrinology*, 8(115):197–202.
- Kim, H. (2005). Parental Investment between Children with Different Abilities. Unpublished manuscript, Department of Economics, University of Wisconsin-Madison.
- Klugman, J., Hanmer, L., Twigg, S., McCleary-Sills, J., Hasan, T., and Bonilla, J. (2014). Voice and Agency: Empowering Women and Girls for Shared Prosperity. Technical report, World Bank.
- Kolbe, A., Hutson, R., Shannon, H., Trzcinski, E., Miles, B., Levitz, N., Puccio, M., James, L., Noel, J., and Muggah, R. (2010). Mortality, crime and access to basic needs before and after the Haiti earthquake: a random survey of Port-au-Prince households. *Medicine, Conflict and Survival*, 26(4):281–297.
- Larrañaga, O. and Herrera, R. (2010a). Efectos en la Calidad de Vida de la Población Afectada por el Terremoto y el Tsunami. Technical report, Ministerio de Planificación.
- Larrañaga, O. and Herrera, R. (2010b). Encuesta Post Terremoto: Principales Resultados. Technical report, Ministerio de Planificación de Chile.
- Lee, D. S. and Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, 48(2):281–355.
- Leight, J. (2014). Sibling Rivalry: Endowment and Intrahousehold Allocation in Gansu Province, China. Unpublished manuscript.
- Leitner, M., Barnett, M., Kent, J., and Barnett, T. (2011). The Impact of Hurricane Katrina on Reported Crimes in Louisiana: A Spatial and Temporal Analysis. *The Professional Geographer*, 63(2):244–261.
- Leitner, M. and Helbich, M. (2011). The Impact of Hurricanes on Crime: A Spatio-Temporal Analysis in the City of Houston, Texas. *Cartography and Geographic Information Science*, 38(2):213–221.

- MacBeath, J. (2010). Living with the Colonial Legacy: The Ghana Story. CCE Report No. 3 3, The Centre for Commonwealth Education.
- Mackie, G., Moneti, F., Shakya, H., and Denny, E. (2015). What are Social Norms? How are They Measured? Technical report, UNICEF.
- Majlesi, K. (2014). Labor Market Opportunities and Women's Decision Making Power within Households. Working Papers 2014:4, Lund University, Department of Economics.
- McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714.
- McGarry, K. and Schoeni, R. F. (1995). Transfer Behavior in the Health and Retirement Study: Measurement and the Redistribution of Resources within the Family. *The Journal of Human Resources*, 30:S184–S226.
- McGovern, M. E. and Canning, D. (2015). Vaccination and all-cause child mortality from 1985 to 2011: global evidence from the Demographic and Health Surveys. *Journal of American Epidemiology*, 182(9):791–798.
- Mechoulan, S. and Wolff, F.-C. (2015). Intra-Household Allocation of Family Resources and Birth Order: Evidence from France using Siblings Data. *Journal of Population Economics*, 28(4):937–964.
- Moghadam, V. (2004). Patriarchy in Transition: Women and the Changing Family in the Middle East. *Journal of Comparative Family Studies*, 35(2):137–162.
- Nahuelpan, E. and Varas, J. (2011). El Terremoto Tsunami en Chile. Una mirada a las estadísticas médico legales. Technical report, Servicio Médico Legal.
- Nguyen, M. C. and Wodon, Q. (2015). *Impact of Early Marriage on Literacy and Educational Attainment in Africa in Child Marriage and Education in Sub-Saharan Africa*. World Bank.
- Olausson, P., Cnattingius, S., and Haglund, B. (1999). Teenage Pregnancies and Risk of Late Fetal Death and Infant Mortality. *British Journal of Obstetric Gynaecology*, 106(2):116–121.

- Olausson, P., Cnattingius, S., and Haglund, B. (2007). Teenage Pregnancies and Adverse Birth Outcomes: A Large Population Based Retrospective Study. *International Journal of Epidemiology*, 36(2):368–373.
- OPM (2010). El Terremoto y Tsunami del 27 de Febrero en Chile. Technical report, Organización Panamericana de la Salud.
- Ormeño, H. (2010). Entendiendo el Comportamiento Cívico Post-Terremoto. Technical report, Universidad de Santiago de Chile.
- Parsons, J., Edmeades, J., Kes, A., Petroni, S., Sexton, M., and Wodon, Q. (2015). Economic Impacts of Child Marriage: A Review of the Literature. *The Review of Faith & International Affairs*, 13(3):12–22.
- Paul Miller, Charles Mulvey, N. M. (1995). What Do Twins Studies Reveal About the Economic Returns to Education? A Comparison of Australian and U.S. Findings. *The American Economic Review*, 85(3):586–599.
- Peacock, W. G., Morrow, B., and Gladwin, H. (1997). *Hurricane Andrew: ethnicity, gender and the sociology of disasters*. London: Routledge.
- Pischke, J.-S. (2007). The Impact of Length of the School Year on Student Performance and Earnings: Evidence From the German Short School Years. *Economic Journal*, 117(523):1216–1242.
- Pitt, M. M., Rosenzweig, M. R., and Hassan, M. N. (1990). Productivity, Health, and Inequality in the Intrahousehold Distribution of Food in Low-Income Countries. *American Economic Review*, 80(5):1139–56.
- Pullum, T. W. (2008). An Assessment of the Quality of Data on Health and Nutrition in the DHS Surveys 1993-2003. Technical report.
- Quarantelli, E. (2001). *The Sociology of Panic*. in In N.J. Smelser and P.B. Baltes (eds.) International Encyclopedia of the Social and Behavioral Sciences.
- Quarantelli, E. and Dynes, R. (1970). Property Norms and Looting: Their Patterns in Community Crises. *Phylon*, 31(2):168–182.

- Raj, A., Saggurthi, N., Winter, M., Labonte, A., Decker, M. R., Balaiah, D., and Silverman, J. G. (2010). The effect of maternal child marriage on morbidity and mortality of children under 5 in India: cross sectional study of a nationally representative sample. *BMJ*, 340.
- Raven, J., Matrices, R. P., Hill, M., and Scales, V. (1994). Manual for Raven's progressive matrices and mill hill vocabulary scales. Advanced progressive matrices.
- Roy, S. (2010). The Impact of Natural Disasters on Crime. Working Papers in Economics 10/57, University of Canterbury, Department of Economics and Finance.
- Sekhri, S. and Debnath, S. (2014). Intergenerational Consequences of Early Age Marriages of Girls: Effect on Children's Human Capital. *Journal of Development Studies*, 50(12):1670–1686.
- Shaw, C. and McKay, H. (1942). *Juvenile Delinquency and Urban Areas*. University of Chicago Press.
- Siegel, J., Bourque, L., and Shoaf, K. (1999). Victimization after a Natural Disaster: Social Disorganization or Community Cohesion? *International Journal of Mass Emergencies and Disasters*, 17(3):265–294.
- Soares, R. R. (2004). Development, Crime and Punishment: Accounting for the International Differences in Crime Rates. *Journal of Development Economics*, 73(1):155–184.
- Solanke, B. (2015). Marriage Age, Fertility Behaviour and Women's Empowerment in Nigeria. Technical report, Obafemi Awolowo University.
- Thistlethwaite, D. L. and Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, 51(6):309–317.
- UN (2013). The Millennium Developments Goals Report. Technical report, United Nations.
- UNEP (2011). On Overview of Our Changing Environment. Technical report, United Nations Environment Programme.
- UNESCO (2014). Education for All 2015: Ghana Country Report. Technical report.

- UNFPA (2012). Marrying Too Young. End Child Marriage. Technical report, United Nations Population Fund.
- UNICEF (2014). Ending Child Marriage: Progress and Prospects. Technical report, United Nations Children's Fund.
- UNISDR (2013). Annual Report. Technical report, The United Nations Office for Disaster Risk Reduction.
- USAID (2009). Basic Education in Ghana: Progress and Problems. Technical report, USAID.
- VanLandingham, M. (2009). Murder Rates in New Orleans, La, 2004–2006. *The Professional Geographer*, 97(9):244–261.
- Wachs, T. (2008). Mechanisms Linking Parental Education and Stunting. *The Lancet*, 371(9609).
- Wahhaj, Z. (2015). A Theory of Child Marriage. Studies in Economics 1520, School of Economics, University of Kent.
- Wodon, Q., Nguyen, M. C., and Tsimpo, C. (2016). Child marriage, education, and agency in uganda. *Feminist Economics*, 22(1):54–79.
- Wolff, F.-C. (2006). Microeconomic models of family transfers. volume 1, chapter 13, pages 889–969. Elsevier, 1 edition.
- Woods, D. L., Kishiyama, M., Yund, E., Herron, T., Edwards, B., Poliva, O., Hink, R., and Reed, B. (2011). Improving Digit Span Assessment of Short-Term Verbal Memory. *Journal of Clinical and Experimental Neuropsychology*, 33(1):101–111.
- Yonetani, M. (2012). Global estimates 2011: People displaced by natural hazard-induced disasters. Technical report, UNHCR.
- Zahran, S., O'Connor, T., and Peek, L. (2009). Natural Disasters and Social Order: Modeling Crime Outcomes in Florida. *International Journal of Mass Emergencies and Disasters*, 27(1):26–52.